THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

The structural characterisation of proteins MAL and Sr33 involved in innate immunity

Peter Lavrencic

Bachelor of Science Hons 1

A thesis submitted for the degree of Doctor of Philosophy at

The University of Queensland in 2016

School of Chemistry & Molecular Biosciences

Abstract

The activation of the Toll-like receptors (TLRs) on eukaryotic cell membranes results in the cytosolic TIR domains coming into close proximity for the binding of adaptor proteins to form a scaffold for downstream signalling. MAL/TIRAP (MyD88 (myeloid differentiation primary response gene 88) adaptor-like/TIR domain-containing adaptor protein) is the bridging adaptor required for MyD88-dependent signaling by a number of TLRs including TLR4. Crystal structures of the MAL TIR domain (MAL^{TIR}) have shown that the structure of the TIR domain is strikingly different to a typical TIR domain fold. The MAL^{TIR} crystal structure contains a long AB loop situated between the αA helix and the βB strand, but no BB loop. This structural re-arrangement has significance as the BB loop in other TIR domains has been associated with protein interactions. In addition, four of the seven cysteine residues were oxidised in the crystal structures, participating in two disulfide bonds. To understand the role of these cysteines, comprehensive biochemical tests and structural characterisation of the protein was carried out.

This thesis presents an NMR solution structure of the $MAL^{TIRC116A}$ mutant, which shows a significant structural rearrangement compared to the crystal structures. The solution structure of $MAL^{TIRC116A}$ presented in this thesis shows a typical TIR domain fold with the functionally important long BB-loop. Strikingly, all seven cysteine residues of this structure appear to be reduced and do not form disulfide bonds. This was not surprising given that the cytosol in which the MAL protein resides is highly reducing. In addition, all cysteines are amendable to modification, such as glutathionylation, and have varying sensitivities in physiological redox conditions. Of particular note, the peak intensity of cysteine 91 decreased by 45% when the redox environment was changed from a physiologically reducing -225 mV to an oxidising -190 mV, suggesting that this cysteine may play an important functional role.

Further to this, the data presented here shows that the MAL^{TIR} resides in a dynamic equilibrium that can be shifted between a monomeric and oligomeric state when exposed to different pH, salt, temperature, and concentration. The identification of this reversible oligomerisation is remarkable and provides a possible mechanism for the regulation of this signalling pathway. Following purification, a dimer species can also separated. This species is not in equilibrium with the monomer and is not covalently bound to the monomer.

Similarly to TIR domains, coiled-coil (CC) domains in plants are involved in defence signalling and implicated in homotypic interactions. In this thesis, the NMR solution structure of a fragment, consisting of residues 6–120, of the wheat steam rust protein Sr33 CC domain is presented. The

structure corresponds to a four-helix bundle with a surface-exposed functionally important EDVID motif. This is significantly different to the published crystal structure of the closely related CC domain from the orthologous barley immunity receptor MLA10 that confers resistance to powdery mildew; however, the structures share striking similarity to the more distantly related Rx protein CC domain from potato. Biophysical studies demonstrate that in solution, all three structures are monomeric and adopt a similar fold. Further to this, data shows that the minimal functional unit consists of residues 1–142 and that *in planta*, self-association correlates strongly with immunity-induced cell death. Collectively, this work contributes to the structural and functional understanding of the CC domains in plants.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

None.

Publications included in this thesis

None.

Contributions by others to the thesis

The contributions of others to this thesis are as follows: Professor Bostjan Kobe was involved in the conception and design of this project. Professor Kobe was also involved with critical revision of all chapters of this thesis. Dr. Thomas Ve provided support in data analysis and initial design of the cDNA of the MAL protein. Dr. Mehdi Mobli operated the NMR magnet, assisted in refinement of both $MAL^{TIRC116A}$ and $Sr33^{6\text{-}120}$ structures, assisted in critical analysis and reviewed all chapters in this thesis. Mr. Lachlan Casey operated and analysed small-angel X-ray scattering (SAXS) data, which was collected at the Australian Synchrotron, and provided figures for this thesis. Dr. Amanda Nouwens operated and analysed mass spectrometry data. Dr. Simon Williams, Mr. Adam Bentham, Mrs. Stella Cesari, Dr. Daniel Ericsson, Professor Peter Anderson, Professor Alan Mark and Professor Peter Dodds assisted with data and writing of chapter 4 of this thesis.

Statement of parts of the thesis submitted to qualify for the award of another degree

None

Acknowledgements

This PhD candidature has been a long journey. I would like to thank, my wife Clare, parents Drago and Lucija, and parents-in-law Geoffrey and Christine for their emotional support through-out this time. Thank you to Professor Bostjan Kobe for providing me with the opportunity to undertake this PhD as well as the technical support that he has provided. Thank you to my co-supervisor Dr. Thomas Ve who invested time into my learning about proteins and laboratory skills that I have gained in biochemistry. My co-supervisor Dr. Mehdi Mobli has been an integral part of this journey and has provided me with emotional and technical support. I would like to thank him for the countless hours that he spent teaching me NMR theory and spending time at the magnet, as well as the many questions that he never got sick of answering. Dr. Simon Williams has also provided much support during my PhD candidature. He's openness to scientific and non-scientific discussion has kept me striving during difficult times. I would like to thank the Kobe lab members; Adam, Andrew, Jannah, Jess, Lachlan, Megan, Max, Ross, Shane and Zehn, and others that have come and gone during this time. The working environment you create provides the back-bone of this successful laboratory; like the amino-acids of a protein. I would also like to thank the Professor Alan Mark and his group for 'adopting' me as a member of their group. I would like to thank David Poger, Bertrand Caron and Michael Corbett for their support, and openness to help. The social and scientific interactions with people from both the Kobe and the Mark group have kept me on path, especially when the journey became challenging. In addition, I would like to thank Amanda Nouwens, Ben Schulz, Kate Stacy and Lachlan Casey for their technical help and support. I have undertaken this thesis, but it would not be complete without the emotional and technical support that has been provided by these kind people around me.

Keywords

Toll/interleukin-1 receptor domain, TIR domain, MAL, Sr33, NMR spectroscopy, Crystallography.

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 060112 Structural Biology	80%
ANZSRC code: 110707 Innate immunity	20%

Fields of Research (FoR) Classification

FoR code: 0601 Biochemistry and Cell Biology	60%
FoR code: 0601 Innate immunity	20%
FoR code: 0607 Plant Biology	20%

Table of contents

List of Figures

List of Tables

Abbreviations

CC	Coiled-coil domain
DTT	1,4-dithiothreitol
GSH	ɣ-L-glutamyl-L-cysteine-glycine
GST	*S*-glutathione transferase
HSQC	Heteronuclear single quantum coherence spectroscopy
LPS	Lipopolysaccharide
MAL	MyD88 adaptor-like
MALS	Multi-angle laser light scattering
MyD88	Myeloid differentiation primary response gene 88
NF-κB	Nuclear factor kappa B
NMR	Nuclear magnetic resonance
NLR	Nucleotide-binding oligomerization domain (NOD)-like receptor
NOD	Nucleotide-binding oligomerization domain
NOX	NADPH-oxidase
PIP-2	phosphatidylinositol-4,5-bisphosphate
ROS	Reactive oxygen species
SAXS	Small angle X-ray scatering
TIR	Toll/interleikin-1 receptor
TIRAP	TIR domain-containing adaptor protein
TLR	Toll-like receptor
TRAF6	TNF receptor-associated factor 6

Preface

The format of this thesis is somewhat different from a typical thesis format, largely because of the broad scope of biological areas covered by the work in the thesis, although throughout focusing on the structural biology of innate immune proteins using NMR spectroscopy. Chapter 1 contains a manuscript that reviews the evidence of redox processes regulating the human innate immune signalling pathways. This provides the background to one part of the thesis, which focuses on the human innate immunity protein MAL and the role of its redox-sensitive cysteine residues. To introduce other areas not covered in Chapter 1, each of the subsequent chapters contains a separate mini-introduction. Similarly, the experimental procedures are described in subsequent chapters, rather than gathered into a separate material and methods chapter. Chapter 2 discusses the biochemical properties of the MAL protein. The methods used in this chapter are interwoven with the results. Part 1 of this chapter outlines the purification of the MAL protein and the redox-sensitive properties of the cysteine residues. This study is taken further and discussed in Chapter 3. Part 2 of Chapter 2 characterizes the oligomeric properties of MAL. As such, this section contains a discussion following these results. Chapter 3 is a manuscript in preparation and contains the most striking results regarding the structure of the MAL protein, including its solution three-dimensional structure. This chapter is divided into manuscript sections that include the introduction, methods, results, discussion and a concluding paragraph. Chapter 4 corresponds to a manuscript in preparation describing the structure and biophysical properties of the plant innate immunity protein Sr33 coiled-coil domain. This chapter corresponds to collaborative work and includes my NMR structural work as the key contribution. Finally, the last chapter entails the conclusions and future directions of the work covered in the thesis. Included in this chapter are some preliminary studies that have been undertaken as a foundation for future work.

Chapter 1

Introduction

1.1 Contribution

This manuscript is a review, which describes the involvement of redox in the human TLR4 innate immune pathways. For this review, I researched the literature and wrote the review. This was reviewed and revised by Dr. Mehdi Mobli and Professor Bostjan Kobe.

Bound to be together: redox control and the TLR4 innate immunity pathways

Peter Lavrencic, Mehdi Mobli, Bostjan Kobe

1.2 Abstract

The innate immunity system, which includes the Toll-like receptor (TLR) pathways, is an important cellular danger-sensing mechanism. Activation of the human TLR4 pathways stimulates the production of highly oxidising species that alter the reducing cytosol. The population of reduced glutathione is decreased, resulting in the oxidation of reduced thiols in the cytoplasm. The cysteine residues on cytosolic proteins are prone to this oxidation and undergo modification by small redox molecules that include glutathione and reactive oxygen species (ROS). These modifications of the signalling proteins in the TLR4 pathway have been shown to behave as a switch that regulates their function. The mechanism of cysteine oxidation and reduction in peroxiredoxin proteins is well understood, but this is not the case for other affected TLR signalling proteins. This review explores the effect that glutathione and ROS have on the human TLR4 innate immune pathway.

1.3 Background

The mammalian cell cytosol is maintained in a reduced state by a high concentration of reduced glutathione. Strong oxidising molecules such as the reactive oxygen species (ROS) are commonly obtained from incomplete metabolism of oxygen during cellular respiration and result in the damage and disruption of lipids, proteins and nucleotides. ROS results in the modification and oxidation of free thiols in the cytoplasm, including the oxidation of the highly abundant reduced glutathione molecules as well as cysteine residues on cytoplasmic proteins. Most of these modifications are reversible and give cysteine residues the ability to act like specific and sensitive switches. Such modifications have been associated with signalling that stimulates cellular defence, proliferation, differentiation and apoptosis. Lipopolysaccharide (LPS) activation of the human Toll-like receptor

(TLR) 4 innate immune pathway has been shown to protect the cell from oxidative stress (1). A number of signalling proteins in the TLR4 pathway have been identified as specific targets for redox modifications. These protein targets are critical for the transduction of the signal that leads to the stimulation of pro-inflammatory cytokine transcription factor NF-κB. In this review, we first aim to address the role of glutathione as a cytosolic redox regulator, followed by the modifications that can occur on cysteine residues of proteins by either glutathione or ROS. Finally, we discuss specific protein signalling targets in the human TLR4 pathway that have been identified to result in signalling in the presence of oxidation.

1.4 Glutathione: a cytosolic redox buffer

Organic thiol-containing small molecules contribute to the highly reducing cytosolic environment of a cell. This reducing potential is determined by a far higher concentration of a reduced moiety (RSH) than the oxidised one (RSSR). The small molecule glutathione is found in a high cytosolic concentration of up to 11 mM, and hence is the major contributor to the reducing potential (2). The eukaryotic cytosol maintains a high reducing environment, with a reduced-to-oxidised glutathione ratio of up to 400:1 (3). The reduced glutathione (GSH) is a tri-peptide (ɤ-L-glutamyl-L-cysteine-glycine) that contains a thiol that can be oxidised. The glutathione redox pair is characterised by the reduced form (GSH), which contains a free thiol, and the oxidised form (GSSG) that connects two GSH molecules by a disulfide bond. This oxidised form can be reversed to the reduced GSH by the enzyme GSH reductase (4). The redox potential of the glutathione redox pair in the cell cytosol is around -240 mV in proliferating cells, around -180 mV in differentiating cells and -160 mV during apoptosis (5). During an immune response, the cell is under stress and contains a higher presence of oxidative species that increase the redox potential. Biochemically, glutathione is a relatively inert molecule; however, with its ability to form the reduced or oxidised species, the molecule has been shown to have protective anti-oxidant abilities by modifying cysteine residues on proteins or detoxifying reactive molecules. The characteristic of GSH as a nucleophilic reductant allows the molecule to capture reactive oxygen and nitrogen species that would otherwise lead to harmful effects including protein aggregation, which characterises the basis of neurodegenerative disorders such as Alzheimer's and Parkinsons's disease (6).

1.5 Cysteine residues are tuneable molecular switches

Most cysteine residues in cytoplasmic proteins have free thiols due to the high concentration and reducing nature of GSH. The polarisable sulfur atom provides a platform for diverse functionality and highly tuneable reactivity, which is controlled by its solvent accessibility and chemical environment that affect its pK_a, and the oxidation state of the sulfur atom. Cysteines with a low pK_a are far more susceptible to post-translational modifications, which are commonly caused by glutathione (glutathionylation) or by ROS (Figure 1.1). Studies have shown that decreasing the GSH/GSSG ratio or adding the ROS molecule hydrogen peroxide at physiologically relevant concentrations (10–400 μM) promotes disulfide-bond formation in proteins that are present in the cytosol. In addition, lower GSH/GSSG ratios increase the extent of glutathionylation of cytosolic proteins, suggesting that the glutathionylation of proteins could be protecting against irreversible hyper-oxidation (7).

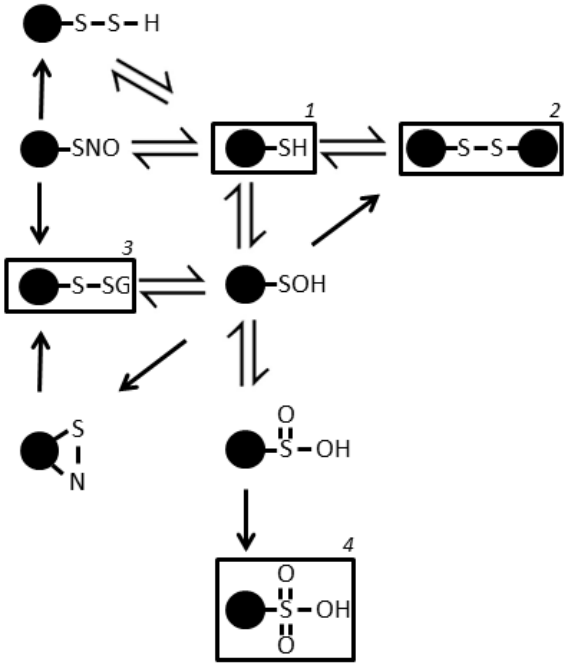


Figure 1.1 Summary of the common redox reactions of cysteine residues in the cytoplasm. *1*: Reactive thiol that is the key to cysteine function. *2*: Disulfide bond formation of two cysteine residues. *3:* Glutathionylation of a cysteine. *4*: Irreversible hyper-oxidation (sulfonic acid).

1.5.1 *Glutathionylation*

A number of studies have suggested that protein glutathionylation can act as a rapid reversible switch that can regulate signalling pathways, including the human TLR4 immune pathway, in a manner that is analogous to phosphorylation (8). The functional role of glutathionylation is

dependent on the transfer of GSH to and from cytosolic proteins by the glutathione *S*-transferase (GST) and glutaredoxin proteins. Specifically, the glutathione redox switch has been identified in the cell cytoplasm as an important regulator in the formation of the cytoskeleton (9), the JNK apoptotic pathway (10), as well as key human transcription factor NF-κB regulating cytokine production (11), which we will focus on in this review.

NF-κB is a protein complex that controls the transcription of a broad range of pro-inflammatory genes (12). Before activation, a family of inhibitors (IκBs) sequester the transcription factor in the cytoplasm. Activation occurs through the IKK (IκB kinase) complex phosphorylating the inhibitory IκB protein, allowing the translocation of NF-κB into the nucleus. Hydrogen peroxide-induced glutathionylation of nuclear factor IKK-β prevents the activation of NF-κB. The oxidation of IKK-β is reversible. The overexpression of glutaredoxin, which reduces the key cysteine 179 by removing glutathione, activates IKK-β even in the presence of hydrogen peroxide (11). Similarly, glutathionylation of other upstream proteins including AKT (13), PTEN (14) and RAS (15) has been shown to represent alternative mechanisms for inhibiting NF-κB signalling.

The GST variant omega-1-1 (GSTO1-1) is an essential protein involved in the transfer of glutathione to and from the TLR4 pathway proteins. The actions of the GSTO1-1 have been shown to generate cytosolic ROS, following the activation of the TLR4 pro-inflammatory pathway (16). Furthermore, GSTO1-1-deficient macrophages are not responsive to LPS and do not activate NF-κB and the NADPH-oxidase NOX1 following TLR4 activation (17). This suggests that the target substrate of GSTO1-1 is likely to be in the TLR adaptor protein complex downstream of TLR4 (Figure 1.2). Knocking out GSTO1-1 results in the increase of GSH and increase in the total protein that has been glutathionylated in macrophage-like cells (17). Unexpectedly, GSTO1-1 is associated with glutathionylation when epithelial breast cancer cells are exposed to high levels of S-nitrosoglutathione (GSNO), followed by subsequent deglutathionlylation catalysis once GSNO is depleted (18). Thus, GSTO1-1 deglutathionylates under confluent cellular conditions, but reverts to glutathionylation under oxidative stress conditions.

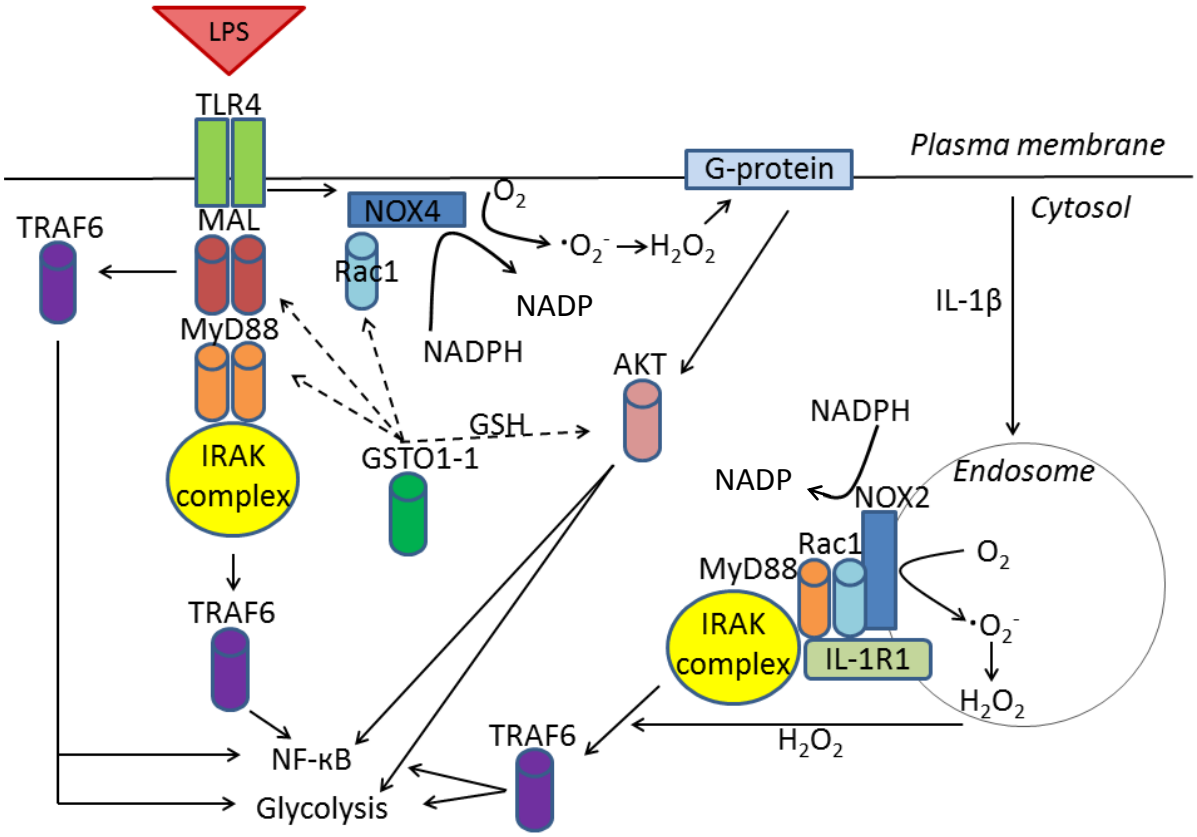


Figure 1.2 The role of ROS and glutathione in the TLR4 signal transduction pathway for the activation of the pro-inflammatory cytokine transcription factor NF-κB. (→) indicates the general activation pathway, (↯) indicates redox reactions, and (-→) indicates likely glutathionylation by the GSTO1-1 protein.

1.5.2 *ROS molecules*

ROS molecules are often produced in eukaryotic cells by the incomplete metabolism of molecular oxygen (O_2) during aerobic respiration and include superoxide ($\cdot O_2^-$), hydrogen peroxide (H_2O_2) and hydroxide radical ($\cdot OH$) (19). These molecules are often dangerous to the cell, resulting in the oxidation of lipids, proteins and DNA, hindering their ability to function (20, 21). Unexpectedly, LPS-induced stimulation of the human immune TLR4 pathway initiates the production of superoxide from NOX proteins through the oxidation of NADPH to $NADP^+$ (discussed later). Superoxide is a highly unstable molecule with an inability to diffuse through the membrane due to its electronegative charge (19-21). Similarly, the hydroxyl radical has a highly indiscriminate reactivity with a half-life of only 10^{-9} s, limiting the molecule to the site of production. Superoxide undergoes both spontaneous and superoxide-dismutase-mediated oxidation to form hydrogen peroxide (19). In contrast to superoxide, hydrogen peroxide is a relatively poor oxidant that reacts slowly with loosely

bound metals (10^3–10^4 $M^{-1}.s^{-1}$), glutathione (2–10 $M^{-1}.s^{-1}$) and reduced thiols on cysteine residues at a neutral pH (19-21). Given the right protein environment, such as a reactive cysteine thiol, the reactivity of hydrogen peroxide to cytosolic protein cysteine residues can increase to 10^7 $M^{-1}.s^{-1}$. This, together with its ability to diffuse through the membrane, makes hydrogen peroxide a prime candidate as a useful signalling ROS molecule in the TLR4 pathway (19-21).

1.5.3 *ROS interaction with peroxiredoxins*

Peroxiredoxins are a well-studied example of a class of anti-oxidant proteins with finely tuned cysteine residues that are highly sensitive to ROS. Peroxiredoxin proteins are highly efficient peroxide-consuming proteins that reduce ROS, while oxidising their own cysteine residues at a rate of 10^7 $M^{-1}.s^{-1}$ (22-24). They have a conserved active-site fold that undergoes a conformational change involving an unfolding event following its activation. The oxidation of the active site cysteine brings two cysteine residues, which are initially around 14 Å apart, into close proximity. The enzymatic mechanism relies on a conserved cysteine residue in the active site that undergoes oxidation from a thiol to a sulfenic acid (SOH) intermediate, which then reacts to form a disulfide bond with the thiol of the second cysteine (25) (Figure 1.3). Oxidation of peroxiredoxin proteins further leads to the formation of larger oligomeric complexes, which have been shown to stimulate the TLR4 pathway to induce inflammatory cytokines exogenously (26, 27). Peroxiredoxins shift between a stable homodimer state and a doughnut-shaped decamer that is destabilised by disulfide bond formation, due to the local unfolding of the structure surrounding the cysteine residues, resulting in a re-modelling of the interacting interface (28). It has been suggested that the decamer is oxidised more efficiently (29), whereas dimers are more effectively reduced by thioredoxins (30). Recycling of the peroxiredoxin from the disulfide-bonded state into the two-free-thiol state is the last step in the cycle. The oxidised thiols are reduced by reacting with a thioredoxin protein, or two reduced glutathione molecules (GSH) to make GSSG (4, 31).

The active cysteine residue C51 of human peroxiredoxin-1 has been shown to hyper-oxidise in sub-millimolar concentrations of hydrogen peroxide. This oxidation occurs at a rate of 0.072% and converts sulfenic acid (SOH) to sulfinic acid (SO_2H), with a minority irreversibly oxidising further to the sulfonic acid (SO_3H) (32). Until recently, the progression to sulfinic acid has been thought to be irreversible; however, the discovery of sulfiredoxin proteins has shown that the reaction can be reversed in an ATP-hydrolysis-dependent manner (33). The ability of peroxiredoxin proteins to

recycle and modify their cysteine residues has shown that these proteins play a wider role in endogenous cytoplasmic (34-36) and exogenous TLR4 signalling pathways.

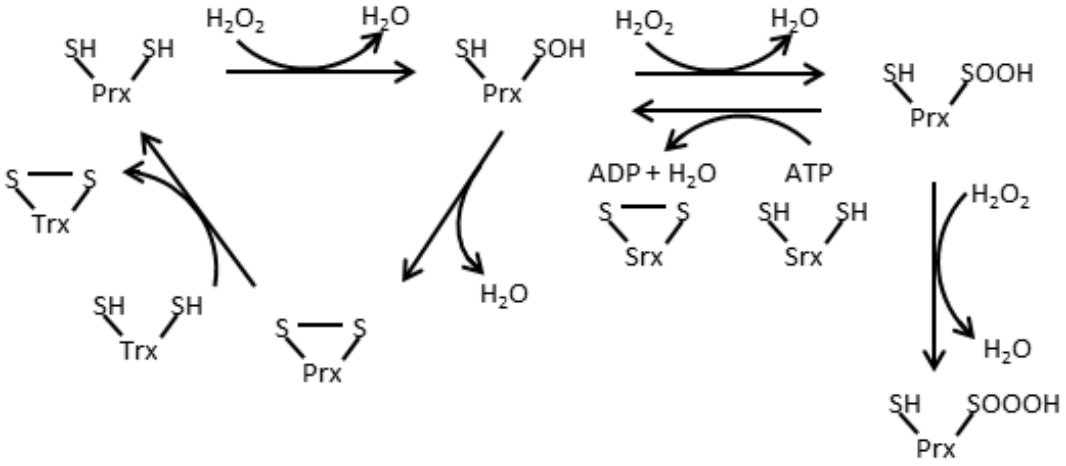


Figure 1.3 Peroxiredoxin redox cycle. Peroxiredoxin (Prx) thiols are oxidised by hydrogen peroxide to form a sulfenic acid intermediate, which usually reacts with a cysteine thiol pair to form a disulfide bond. The disulfide bond can be recycled by a thioreductase (Trx). Oxidation of the sulfenic acid (SOH) can result in a sulfinic acid, which can be reversed by a sulfiredoxin and ATP hydrolysis; however, oxidation of sulfinic acid (SO_2H) into sulfonic acid (SO_3H) is an irreversible reaction (31).

1.6 Redox signalling in the TLR4 pathways

Induction by LPS or exogenous peroxiredoxin protein simulates the association of TLR4 proteins to come into close proximity and produce a platform for a downstream signalling cascade that activates the pro-inflammatory transcription factor NF-κB (37). Here we discuss three major pathways, shown in Figure 1.2, that utilise small redox molecules (ROS and glutathione) for their signalling to induce NF-κB.

1.6.1 *MyD88-dependent pathway*

The MyD88-dependent pathway is initiated by an interaction between the cytoplasmic TIR (Toll/interleikin-1 receptor) domain of TLR4 and the TIR domains from TLR adaptors MAL and MyD88. The structure of MAL is predominantly a TIR domain and acts as a bridging adaptor between the receptor and the downstream protein MyD88 (38). The MyD88 protein consists of two domains, namely, a TIR domain and a death domain. The TIR domain forms TIR-TIR domain interactions with

MAL, while the death domain is involved in Myddosome assembly formation. The crystal structure of this assembly solved by Lin *et al* (39)demonstrates that death domains form the core interface between MyD88, and the IRAK (interleukin-1 receptor-associated kinase) proteins with a ratio of 6:4:4 for MyD88:IRAK4:IRAK2, respectively (39). MyD88 has been associated with all cell-surface TLRs (2, 4, and 5) as well as endosomal TLRs 7, 8, and 9, indicating that this is a key molecule in the TLR signalling pathway (38). The recruitment of downstream protein TRAF6 (TNF receptor-associated factor 6) has been shown to occur by either direct association with MAL, or by the interaction with the IRAK complex (Figure 1.2) (40), however, it is unclear what determines this selection, or activates signal transduction. A closer examination of the crystal structure of MAL TIR domain suggests a possible redox mechanism that could drive this pathway. The crystal structure of the MAL TIR domain shows cysteine residues forming two disulfide bonds through cysteines 89-134 and 142-174 (41-44). The highly reducing environment suggests that these cysteine residues are reduced in the cytosol but are prone to oxidation. The regulation of the formation of the TIR-domain assembly is unclear; however, recent work by Menon *et al*. (17) suggests that the glutathione transferase GSTO1-1 protein targets these proteins and plays a role in the transfer of glutathione, which act as a switch (17).

1.6.2 NOX4-dependent pathway

NOX proteins have been widely studied in phagocytic cells for their role in the generation of superoxide and hydrogen peroxide (45). The discovery of several classes of NOX proteins has found that they are specific to cell types and may possess unique functions (46). A study by Park *et al* (47) found that the C-terminal region of the NOX4 protein interacts directly with the cytosolic TIR domain of TLR4 upon LPS activation, to stimulate the production of ROS. Another recent study supports this and suggests that MyD88 is critical in the activation of NOX proteins, which leads to the production of ROS molecules (48). Together, these studies suggest that like the IRAKs, NOX4 also interacts with the TLR4-MAL-MyD88 TIR domain scaffold. ROS production has been reported to increase upon LPS stimulation, which is a result of the NOX4 protein catalysing the production of superoxide and simultaneously forming hydrogen peroxide from molecular oxygen, following the oxidation of NADPH (47, 49). Hydrogen peroxide activates the PI3 kinase (phosphatidylinositol-3-kinase), which transduces the signal to a G-protein, followed by AKT activation, pro-inflammatory gene expression and glucose uptake (50) (Figure 1.3). *In vivo* data has demonstrated that knocking out the NOX4 protein results in a 30% decrease of NF-κB activation and no activation of AKT following LPS-

treatment, suggesting that the production of ROS via the NOX4-dependent pathway is important for TLR4 signalling (47).

1.6.3 Endosomal pathway

Pioneering work by Li *et al* in 2005, described a mechanism by which NOX2 provides the redox signalling molecules for the activation of NF-κB. Results from this study demonstrated that upon the activation of the interleukin-1 receptor, interleukin-1β recruits MyD88 and the small GTPase Rac1 at the plasma membrane. *In vivo* experiments demonstrated that Rac1 was required for the recruitment of NOX2 to the endosome, while MyD88 induced the formation of the endosome. This resulted in the activation of the NOX2 protein in the endosome for the production of ROS molecules, which activated the recruitment of the IRAK complex and TRAF6 to the receptor (51). This pathway represents a subcellular framework for spatially controlled ROS production by NOX2. The recruitment of TRAF6 to the receptor is directly dependent on hydrogen peroxide. The mechanism of TRAF6 recruitment is unclear, but possibly results in a redox-dependent modification of either TRAF6 or an upstream protein (51).

1.7 Concluding remarks

Small redox molecules, including glutathione and ROS, play an important role in cell signalling. The versatility of cysteine residues is protein-specific and relates to the specific protein function. The importance of cysteine reactivity in TLR4 signalling proteins has been identified, however, the mechanism in many cases remains elusive. Proteins such as peroxiredoxins provide a mechanistic insight into the diversity of oxidising modifications and structural changes that take part during an oxidation event. Most of the oxidising modifications, such as glutathionylation, disulfide-bond formation and sulfinic acid formation are reversible and are utilised for activation or inhibition of protein signalling. This review has focussed on the importance of redox molecules in the human TLR4 immune signalling pathway, which is just one of many systems these molecules affect. In this and other pathways, there remain many gaps in our understanding of the role of small redox molecules that were once thought to be only a hazard to the human cell.

1.8 Chapter 1 references

1. Deng, S., Yu, K., Wu, Q., Li, Y., Zhang, X., Zhang, B., Liu, G., Liu, Y., and Lian, Z. (2015) Toll-Like Receptor 4 reduces oxidative injury via glutathione activity in sheep, *Oxidative Medicine and Cellular Longev*ity. Article: 9151290.
2. Smith, C. V., Jones, D. P., Guenthner, T. M., Lash, L. H., and Lauterburg, B. H. (1996) Compartmentation of Glutathione: Implications for the Study of Toxicity and Disease, *Toxicology and Applied Pharmacology* **140**, 1–12.
3. Gilbert, H. F. (1995) Thiol/disulfide exchange equilibria and disulfide bond stability, *Methods in Enzymology* **251**, 8–28.
4. Aquilano, K., Baldelli, S., and Ciriolo, M. R. (2014) Glutathione: new roles in redox signaling for an old antioxidant, *Frontiers in Pharmacology* **5**, 196.
5. Sarsour, E. H., Kumar, M. G., Chaudhuri, L., Kalen, A. L., and Goswami, P. C. (2009) Redox control of the cell cycle in health and disease, *Antioxidants & Redox Signaling* **11**, 2985-3011.
6. Chung, K. K. (2006) Say NO to Neurodegeneration: Role of S-Nitrosylation in Neurodegenerative Disorders, *Neurosignals* **15**, 307–313.
7. Cumming, R. C., Andon, N. L., Haynes, P. A., Park, M., Fischer, W. H., and Schubert, D. (2004) Protein disulfide bond formation in the cytoplasm during oxidative stress, *The Journal of Biological Chemistry* **279**, 21749–21758.
8. Ellis, J. J., and Kobe, B. (2013) Phosphorylation., In *Encyclopedia of Biophysics* (Roberts, G. C. K., Ed.), 559–563, Springer-Verlag, Berlin Heidelberg.
9. Fiaschi, T., Cozzi, G., Raugei, G., Formigli, L., Ramponi, G., and Chiarugi, P. (2006) Redox regulation of beta-actin during integrin-mediated cell adhesion, *The Journal of Biological Chemistry* **281**, 22983–22991.
10. Townsend, D. M., Manevich, Y., He, L., Hutchens, S., Pazoles, C. J., and Tew, K. D. (2009) Novel role for glutathione s-transferase π: regulation of protein s-glutathionylation following oxidative and nitrosative stress, *The Journal of Biological Chemistry* **284**, 436–445.
11. Reynaert, N. L., van der Vliet, A., Guala, A. S., McGovern, T., Hristova, M., Pantano, C., Heintz, N. H., Heim, J., Ho, Y. S., Matthews, D. E., Wouters, E. F., and Janssen-Heininger, Y. M. (2006) Dynamic redox control of NF-kappaB through glutaredoxin-regulated S-glutathionylation of inhibitory kappaB kinase beta, *Proceedings of the National Academy of Sciences U S A* **103**, 13086–13091.
12. Hacker, H., and Karin, M. (2006) Regulation and function of IKK and IKK-related kinases, *Science Signaling 2006*, re13.

13. Murata, H., Ihara, Y., Nakamura, H., Yodoi, J., Sumikawa, K., and Kondo, T. (2003) Glutaredoxin exerts an antiapoptotic effect by regulating the redox state of Akt, *The Journal of Biological Chemistry* **278**, 50226–50233.

14. Cruz, C. M., Rinna, A., Forman, H. J., Ventura, A. L., Persechini, P. M., and Ojcius, D. M. (2007) ATP activates a reactive oxygen species-dependent oxidative stress response and secretion of proinflammatory cytokines in macrophages, *The Journal of Biological Chemistry* **282**, 2871–2879.

15. Clavreul, N., Bachschmid, M. M., Hou, X., Shi, C., Idrizovic, A., Ido, Y., Pimentel, D., and Cohen, R. A. (2006) S-glutathiolation of p21ras by peroxynitrite mediates endothelial insulin resistance caused by oxidized low-density lipoprotein, *Arteriosclerosis, Thrombosis,and Vascular Biology* **26**, 2454–2461.

16. Menon, D., Coll, R., O'Neill, L. A. J., and Board, P. G. (2014) Glutathione transferase Omega 1 is required for the lipopolysaccharide-stimulated induction of NADPH oxidase 1 and the production of reactive oxygen species in macrophages, *Free Radical Biology & Medicine* **73**, 318–327.

17. Menon, D., Coll, R., O'Neill, L. A. J., and Board, P. G. (2015) GSTO1-1 modulates metabolism in macrophages activated through the LPS and TLR4 pathway, *Journal of Cell Science* **128**, 1982–1990.

18. Menon, D., and Board, P. G. (2013) A Role for Glutathione Transferase Omega 1 (GSTO1-1) in the glutathionylation cycle, *The Journal of Biological Chemistry* **288**, 25769–25779.

19. Imlay, J. A. (2003) Pathways of Oxidative Damage, *Annual Review of Microbiology* **57**, 395–418.

20. Winterbourn, C. C., and Metodiewa, D. (1999) Reactivity of biologically important thiol compounds with superoxide and hydrogen peroxide, *Free Radical Biology & Medicine* **27**, 322–328.

21. D'Autreaux, B., and Toledano, M. B. (2007) ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis, *Nature Reviews Molecular Cell Biology* **8**, 813–824.

22. Winterbourn, C. C. (2008) Reconciling the chemistry and biology of reactive oxygen species, *Nature Chemical Biology* **4**, 278–286.

23. Ogusucu, R., Rettori, D., Munhoz, D. C., Netto, L. E., and Augusto, O. (2007) Reactions of yeast thioredoxin peroxidases I and II with hydrogen peroxide and peroxynitrite: rate constants by competitive kinetics, *Free Radical Biology & Medicine* **42**, 326–334.

24. Peskin, A. V., Low, F. M., Paton, L. N., Maghzal, G. J., Hampton, M. B., and Winterbourn, C. C. (2007) The high reactivity of peroxiredoxin 2 with H(2)O(2) is not reflected in its reaction

with other oxidants and thiol reagents, *The Journal of Biological Chemistry* **282**, 11885–11892.

25. Nelson, K. J., Parsonage, D., Hall, A., Karplus, P. A., and Poole, L. B. (2008) Cysteine pK(a) values for the bacterial peroxiredoxin AhpC, *Biochemistry* **47**, 12860–12868.
26. Riddell, J. R., Wang, X. Y., Minderman, H., and Gollnick, S. O. (2010) Peroxiredoxin 1 stimulates secretion of proinflammatory cytokines by binding to TLR4, *Journal of Immunology* **184**, 1022–1030.
27. Shichita, T., Hasegawa, E., Kimura, A., Morita, R., Sakaguchi, R., Takada, I., Sekiya, T., Ooboshi, H., Kitazono, T., Yanagawa, T., Ishii, T., Takahashi, H., Mori, S., Nishibori, M., Kuroda, K., Akira, S., Miyake, K., and Yoshimura, A. (2012) Peroxiredoxin family proteins are key initiators of post-ischemic inflammation in the brain, *Nature Medicine* **18**, 911–917.
28. Wood, Z. A., Poole, L. B., Hantgan, R. R., and Karplus, P. A. (2002) Dimers to doughnuts: redox-sensitive oligomerization of 2-cysteine peroxiredoxins, *Biochemistry* **41**, 5493–5504.
29. Parsonage, D., Youngblood, D. S., Sarma, G. N., Wood, Z. A., Karplus, P. A., and Poole, L. B. (2005) Analysis of the link between enzymatic activity and oligomeric state in AhpC, a bacterial peroxiredoxin, *Biochemistry* **44**, 10583–10592.
30. Matsumura, T., Okamoto, K., Iwahara, S., Hori, H., Takahashi, Y., Nishino, T., and Abe, Y. (2008) Dimer-oligomer interconversion of wild-type and mutant rat 2-Cys peroxiredoxin: disulfide formation at dimer-dimer interfaces is not essential for decamerization, *The Journal of Biological Chemistry* **283**, 284–293.
31. Poole, L. B. (2007) The catalytic mechanism of peroxiredoxins, *Subcellular Biochemistry* **44**, 61–81.
32. Yang, K. S., Kang, S. W., Woo, H. A., Hwang, S. C., Chae, H. Z., Kim, K., and Rhee, S. G. (2002) Inactivation of human peroxiredoxin I during catalysis as the result of the oxidation of the catalytic site cysteine to cysteine-sulfinic acid, *The Journal of Biological Chemistry* **277**, 38029–38036.
33. Biteau, B., Labarre, J., and Toledano, M. B. (2003) ATP-dependent reduction of cysteine-sulphinic acid by S. cerevisiae sulphiredoxin, *Nature* **425**, 980–984.
34. Kuge, S., Arita, M., Murayama, A., Maeta, K., Izawa, S., Inoue, Y., and Nomoto, A. (2001) Regulation of the yeast Yap1p nuclear export signal is mediated by redox signal-induced reversible disulfide bond formation, *Molecular and Cellular Biology* **21**, 6139–6150.
35. Toledano, M. B., Delaunay, A., Monceau, L., and Tacnet, F. (2004) Microbial H2O2 sensors as archetypical redox signaling modules, *Trends in Biochemical Sciences* **29**, 351–357.

36. Izawa, S., Maeda, K., Sugiyama, K., Mano, J., Inoue, Y., and Kimura, A. (1999) Thioredoxin deficiency causes the constitutive activation of Yap1, an AP-1-like transcription factor in Saccharomyces cerevisiae, *The Journal of Biological Chemistry* **274**, 28459–28465.

37. Gay, N. J., Symmons, M. F., Gangloff, M., and Bryant, C. E. (2014) Assembly and localization of Toll-like receptor signalling complexes, *Nature Reviews Immunology* **14**, 546–558.

38. Ve, T., Gay, N. J., Mansell, A., Kobe, B., and Kellie, S. (2012) Adaptors in toll-like receptor signaling and their potential as therapeutic targets, *Current Drug Targets* **13**, 1360–1374.

39. Lin, S.-C., Lo, Y.-C., and Wu, H. (2010) Helical assembly in the MyD88–IRAK4–IRAK2 complex in TLR/IL-1R signalling, *Nature* **465**, 885–890.

40. Verstak, B., Nagpal, K., Bottomley, S. P., Golenbock, D. T., Hertzog, P. J., and Mansell, A. (2009) MyD88 Adapter-like (Mal)/TIRAP interaction with TRAF6 is critical for TLR2- and TLR4-mediated NF-κB proinflammatory responses, *The Journal of Biological Chemistry* **284**, 24192–24203.

41. Valkov, E., Stamp, A., DiMaio, F., Baker, D., Verstak, B., Roversi, P., Kellie, S., Sweet, M. J., Mansell, A., Gay, N. J., Martin, J. L., and Kobe, B. (2011) Crystal structure of Toll-like receptor adaptor MAL/TIRAP reveals the molecular basis for signal transduction and disease protection, *Proceedings of the National Academy of Sciences U S A* **108**, 14879–14884.

42. Lin, Z., Lu, J., Zhou, W., and Shen, Y. (2012) Structural insights into TIR domain specificity of the bridging adaptor Mal in TLR4 signaling, *PLoS One* **7**, e34202.

43. Snyder, G. A., Deredge, D., Waldhuber, A., Fresquez, T., Wilkins, D. Z., Smith, P. T., Durr, S., Cirl, C., Jiang, J., Jennings, W., Luchetti, T., Snyder, N., Sundberg, E. J., Wintrode, P., Miethke, T., and Xiao, T. S. (2014) Crystal structures of the Toll/Interleukin-1 receptor (TIR) domains from the Brucella protein TcpB and host adaptor TIRAP reveal mechanisms of molecular mimicry, *The Journal of Biological Chemistry* **289**, 669–679.

44. Woo, J. R., Kim, S., Shoelson, S. E., and Park, S. (2012) X-ray Crystallographic Structure of TIR-Domain from the Human TIR-Domain Containing Adaptor Protein/MyD88-Adaptor-Like Protein (TIRAP/MAL), *Bulletin of the Korean Chemical Society* **33**, 3091–3094.

45. Bokoch, G. M., and Knaus, U. G. (2003) NADPH oxidases: not just for leukocytes anymore!, *Trends in Biochemical Sciences* **28**, 502–508.

46. Lambeth, J. D. (2002) Nox/Duox family of nicotinamide adenine dinucleotide (phosphate) oxidases, *Current Opinions in Hematology* **9**, 11–17.

47. Park, H. S., Jung, H. Y., Park, E. Y., Kim, J., Lee, W. J., and Bae, Y. S. (2004) Cutting Edge: Direct interaction of TLR4 with NAD(P)H oxidase 4 isozyme is essential for lipopolysaccharide-

induced production of reactive oxygen species and activation of NF-κB, *Journal of Immunology* **173**, 3589–3593.

48. Ichikawa, S., Miyake, M., Fujii, R., and Konishi, Y. (2012) MyD88 associated ROS generation is crucial for Lactobacillus induced IL-12 production in macrophage, *PLoS One* **7**, e35880.

49. Ago, T., Kitazono, T., Ooboshi, H., Iyama, T., Han, Y. H., Takada, J., Wakisaka, M., Ibayashi, S., Utsumi, H., and Iida, M. (2004) Nox4 as the major catalytic component of an endothelial NAD(P)H oxidase, *Circulation* **109**, 227–233.

50. Ngkelo, A., Meja, K., Yeadon, M., Adcock, I., and Kirkham, P. A. (2012) LPS induced inflammatory responses in human peripheral blood mononuclear cells is mediated through NOX4 and G(i)α dependent PI-3kinase signalling, *Journal of Inflammation (London)* **9**, 1–1.

51. Li, Q., Harraz, M. M., Zhou, W., Zhang, L. N., Ding, W., Zhang, Y., Eggleston, T., Yeaman, C., Banfi, B., and Engelhardt, J. F. (2006) Nox2 and Rac1 regulate H(2)O(2)-dependent recruitment of TRAF6 to endosomal interleukin-1 receptor complexes, *Molecular Cell Biology* **26**, 140–154.

Chapter 2

The monomeric and oligomeric nature of the MAL TIR domain

2.1 Introduction

The MyD88 adaptor-like (MAL) protein, also known as the Toll/ interleukin-1 receptor (TIR) domain-containing adaptor protein (TIRAP), is a key player in the Toll-like receptor (TLR) signalling cascade of the human immune system. The stimulation of cell surface TLRs results in the cytosolic TIR domains from at least two TLRs to come into close proximity and form a scaffold for downstream signalling (1, 2). The cytosolic MAL protein is recruited upon activation of TLR4 and TLR5, or the combination of TLR2 with either TLR 1 or 6 (3-7) (Figure 2.1).

The MAL adaptor protein contains a phosphatidylinositol-4,5-bisphosphate (PIP-2) binding motif at the N-terminus, which allows the protein to anchor to the plasma membrane (8). The remaining protein (residues 79–221) consists of a TIR domain, which acts as a bridging adaptor between the cytosolic TIR domain of TLR and the MyD88 signalling protein. This set of interactions is dominated by TIR:TIR domain interfaces, which form a platform for downstream signalling proteins such as the IRAKs and TRAF6 that in turn activate NF-κB and the production of pro-inflammatory cytokines (9) (Figure 2.1). Single-nucleotide polymorphisms on the MAL TIR domain have shown to detriment NF-κB activation and compromise an immune response. The rare D96N mutation results in a protein that cannot interact with the downstream MyD88 protein and is thus unable to activate NF-κB (10). Further to this, individuals who are homozygous for the S180L mutation in the MAL TIR domain are more prone to sepsis (11) and do not respond to *Haemophilus influenza* serotype b vaccine (12), while heterozygous individuals have a greater resistance to disease (11). The S180L mutation in heterozygous individuals has been shown to be a protective factor against the development of pneumococcal disease, bacteremia, malaria and, tuberculosis (13, 14). These studies indicate the importance of subtle changes in the early signalling stages of this pathway, leading to significant effects during an immune response.

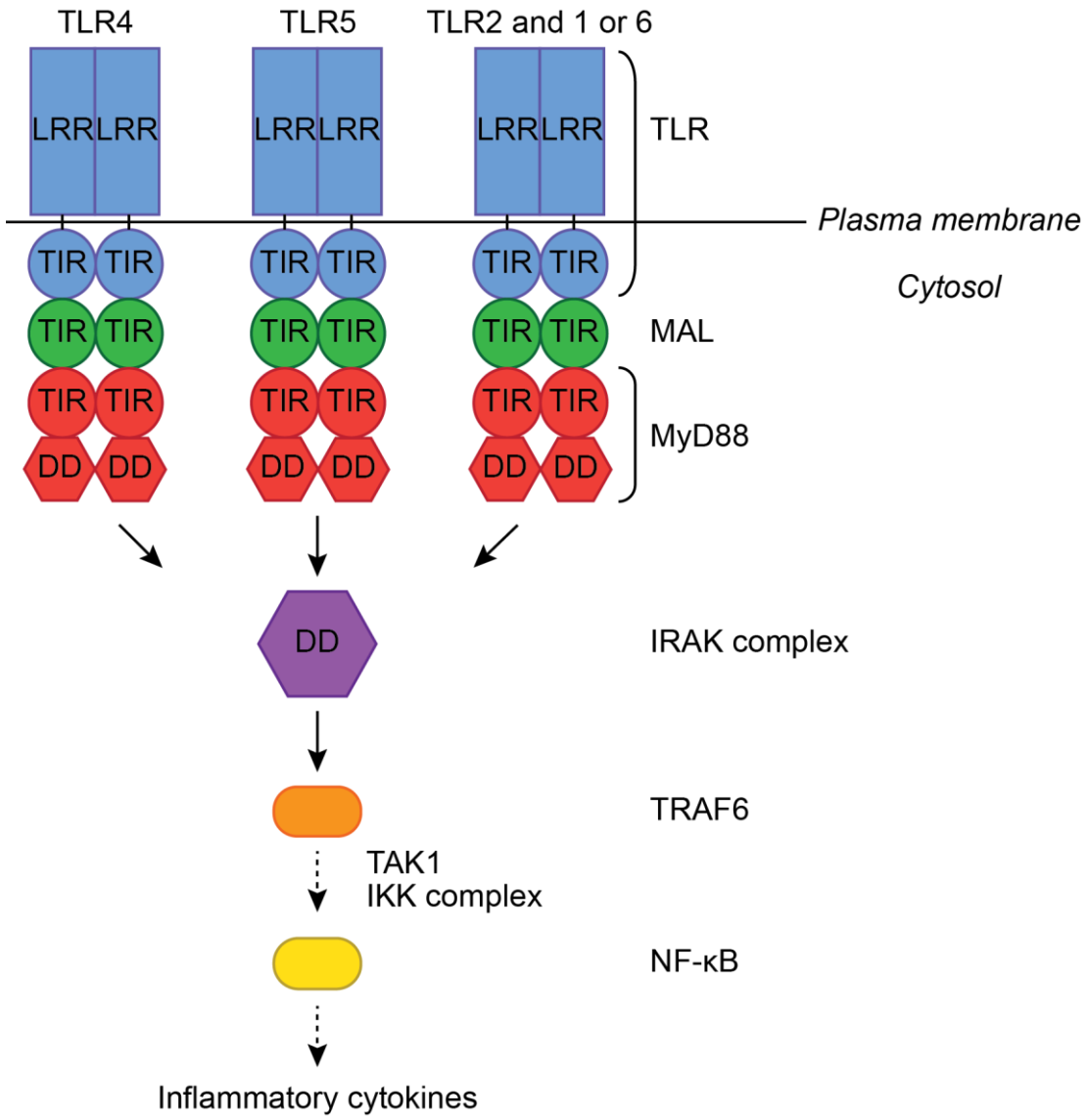


Figure 2.1 MyD88-dependent signalling pathway of Toll-like receptors (TLRs) that are present on the plasma membrane. Two activated TLRs each consisting of an extra-cellular leucine-rich-repeat (LRR) domain and an intracellular TIR domain, form a scaffold for downstream signalling to MAL and MyD88 proteins, which interact using TIR domains. The signal is transduced between MyD88 and IRAK proteins using death-domains (DD).

Four published crystal structures of the MAL TIR domain (MAL^{TIR}) (residues 79–221) protein reveal a mixture of reduced and oxidised cysteine residues. All four crystal structures reported two disulfide bonds, which form between cysteine residues 89:134 and 142:174 (Figure 2.2) (15-18). The MAL protein is localised in the cytosol, which is a reducing environment. Therefore, the presence of disulfide bonds in the cytosol is intriguing and suggests a possible regulatory mechanism for the signalling of the MAL^{TIR} via the redox active cysteine residues.

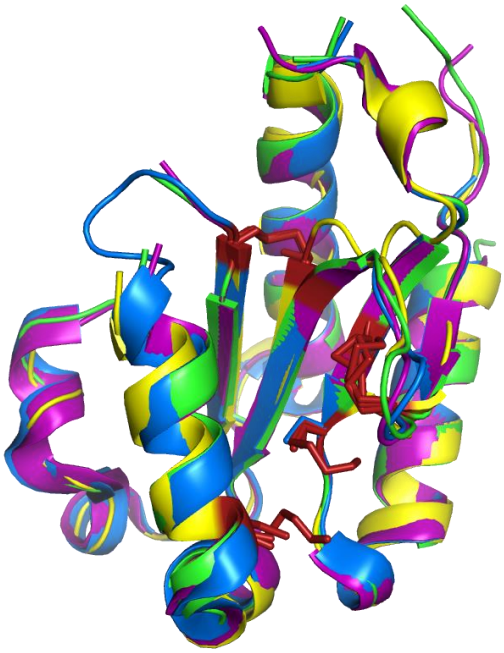


Figure 2.2 Overlay of MAL^{TIR} crystal structures from Valkov 2YP2 (green)(15), Lin 3UB2 (blue)(18), Snyder 4LQD (purple)(17) and Woo 4FZ5 (yellow)(16), showing a mixture of reduced and oxidised cysteine residues in red.

Recent studies have shown that redox molecules, including glutathione and hydrogen peroxide, play an integral part in the activation and the regulation of this pathway. The transfer of glutathione to and from proteins involved in the early stages of the TLR4 signalling cascade is critical for signal transduction following the detection of lipopolysaccharide (LPS). Work by Menon *et al* (19) has demonstrated that knocking out *S*-glutathione transferase (GST) omega class 1 protein, which is responsible for the transfer of glutathione to and from proteins, results in the loss of transcription factor NF-κB activation (19). Similarly, hydrogen peroxide is a critical component that activates the protein TRAF6 downstream (20). On the other hand, the oxidation of cysteine residues on the PTEN phosphatase protein, which is involved in regulating PIP-2 levels and mediating the recruitment of MAL to the cell membrane, results in the protein becoming inactive (21). In addition, this oxidation results in a portion of PTEN forming a higher molecular weight structure that is reversible to monomers under reducing conditions (21).

Results in this chapter reveal that the cysteine residues of monomeric MAL^{TIR} are redox-sensitive. In addition, MAL^{TIR} can be purified as a monomer or a dimer. Interestingly, the two do not appear to be in a dynamic equilibrium on timescales measured here, such that if the dimer is isolated using size-exclusion chromatography and subjected to a subsequent size-exclusion chromatography step, no

monomer is detected. Literature suggests that TIR domains are active as dimers (22), however, the crystal structure was solved using the purified monomer.

This chapter will discuss two aspects of MAL^{TIR}; part 1 will discuss the redox state of the cysteine residues, as well as the dynamic equilibrium between the monomeric and oligomeric states of MAL^{TIR} in solution, while part 2 will discuss the stable dimer of MAL^{TIR}.

2.2 PART 1: The solubility of MAL^{TIR} and redox state of its cysteine residues

2.2.1 *MAL^{TIR} expressed recombinantly in different cell types has the same purification profile*

The gene encoding the human wild-type MAL^{TIR}, consisting of residues 79-221, was cloned into a pMCSG7 vector (23). This region was selected as it had previously been studied and led to four published crystal structures (Figure 2.2). The protein was expressed as a His-fusion protein and purified in three different *E. coli* cell types (BL-21 [New England Biolabs], SHuffle [New England Biolabs] (24), and Origami [Merck Millipore]) using the auto-induction method (25). The protein was purified using nickel affinity chromatography followed by incubation with TEV protease to cleave off the 6-residue histidine tag. The protein was re-eluted over the nickel column and purified further using size-exclusion chromatography [Superdex S75 26/600, GE] into a buffer consisting of 20 mM Tris at pH 8.6 with 200 mM NaCl. BL-21 cells have a reducing cytosolic environment, which is similar to that in eukaryotic cells. On the other hand, SHuffle and Origami are strains that have been engineered with knock-outs to cytoplasmic reductive pathways, resulting in an oxidative cytosolic environment, which promotes the formation of cytoplasmic disulfide bonds (24). Expression of wild-type MAL^{TIR} and subsequent purification under non-reducing conditions resulted in a similar elution profile across all *E. coli* strains tested, when separated by size-exclusion chromatography, as shown in Figure 2.3. The major peak eluted at 200 mL with a smaller shoulder peak at 170 mL and a void volume peak at 120 mL. Both the major elution and shoulder peaks corresponded to the size of the MAL^{TIR} monomer, as confirmed by SDS gel shown in Figure 2.3.

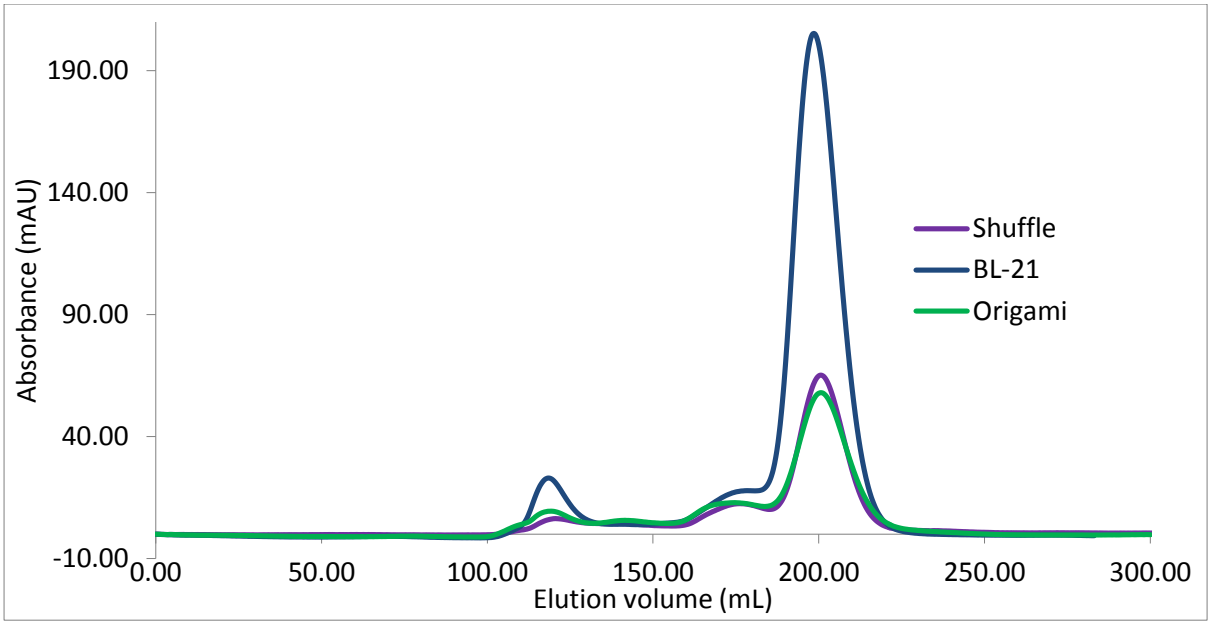


Figure 2.3 Purification curves of MAL^{TIR} using a Superdex 75 26/600 [GE] gel filtration column, following expression and purification in three different cell types.

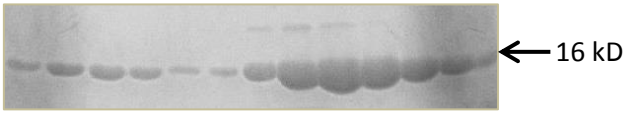


Figure 2.4 Sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) of MAL^{TIR} expression in BL-21 cells following size-exclusion chromatography. Each lane represents a 5 mL fraction between 160 and 220 mL.

2.2.2 *MAL^{TIR} undergoes redox change of cysteine residues when exposed to different redox conditions*

Wild-type MAL^{TIR} was purified from three types of *E. coli* cells (BL-21, SHuffle and Origami), each with different cytosolic redox environments. Ellman's reagent (dithiobisnitrobenzoic acid) (26) was used to analyse the proportion of reduced cysteine residues in MAL^{TIR}. Ellman's reagent reacts with reduced cysteine thiols to form a disulfide bond and produce a yellow-coloured solution (26). The colour became more intense as the number of reduced cysteines reacted with the reagent (26). The colour change in solution was measured at 412 nm using a spectrophotometer [NanoDrop, Thermo Scientific]. The percent of reduced thiols in solution was used to calculate to the nearest number of reduced cysteine residues on MAL^{TIR} (Table 2.1). The analysis revealed that expression of MAL^{TIR} in the reducing cytosolic environment of BL-21 cells resulted, on average, in one oxidised and six reduced cysteine residues. By contrast, expression of MAL^{TIR} in the SHuffle and Origami cell types with a more oxidising cytosolic environment produced three oxidised and four reduced cysteine

residues. Exposing MAL^{TIR} to the strong oxidising agent hydrogen peroxide resulted in the oxidation of all cysteines, while exposure to the reducing agent 1,4-dithiothreitol (DTT) led to the reduction of all cysteine residues. These results indicate that the cysteine residues are sensitive to redox changes. To better understand which of these cysteines were susceptible to redox change in a physiological cell environment, single cysteine mutants were expressed in both the reducing BL-21 and oxidising SHuffle cytosolic cell environments.

Table 2.1 Average number of reduced and oxidised cysteine residues in the MAL^{TIR} protein following expression in different cell types and exposure to chemicals using Elman's assay.

Wild-type MAL^{TIR} expression cell or chemical condition	% Reduced	Reduced	Oxidised
BL-21 with 10 mM DTT and size-exclusion chromatography	100	7	0
BL-21	90	6	1
SHuffle	57	4	3
Origami	53	4	3
H_2O_2 (0.03 mM)	0	0	7

2.2.3 *Purification of MAL^{TIR} single cysteine mutants in BL-21 cells*

All single cysteine-to-alanine mutants were successfully expressed in *E. coli* BL-21 cells (which contain a reducing cytosol), and purified in one litre of LB media using the same protocol as above by nickel affinity and size-exclusion chromatography. Following size-exclusion chromatography, all proteins showed a two peak profile on a gel filtration chromatogram [Superdex 75 26/600, GE] as shown in Figure 2.5. Mutants expressed very different yields of protein, with C89A, C91A, C116A and C134A expressing almost twice as much as C142A, C157A and C174A. The first peak on the chromatogram eluted at around 170 mL, while the second peak eluted at 200 mL, except for mutant C157A, which eluted at 210 mL. The first peak amounted to around 10% of the second peak. Comparatively to other mutants, C89A produced a higher ratio of the first peak compared to second peak, suggesting that this mutation could be important for the formation of this product.

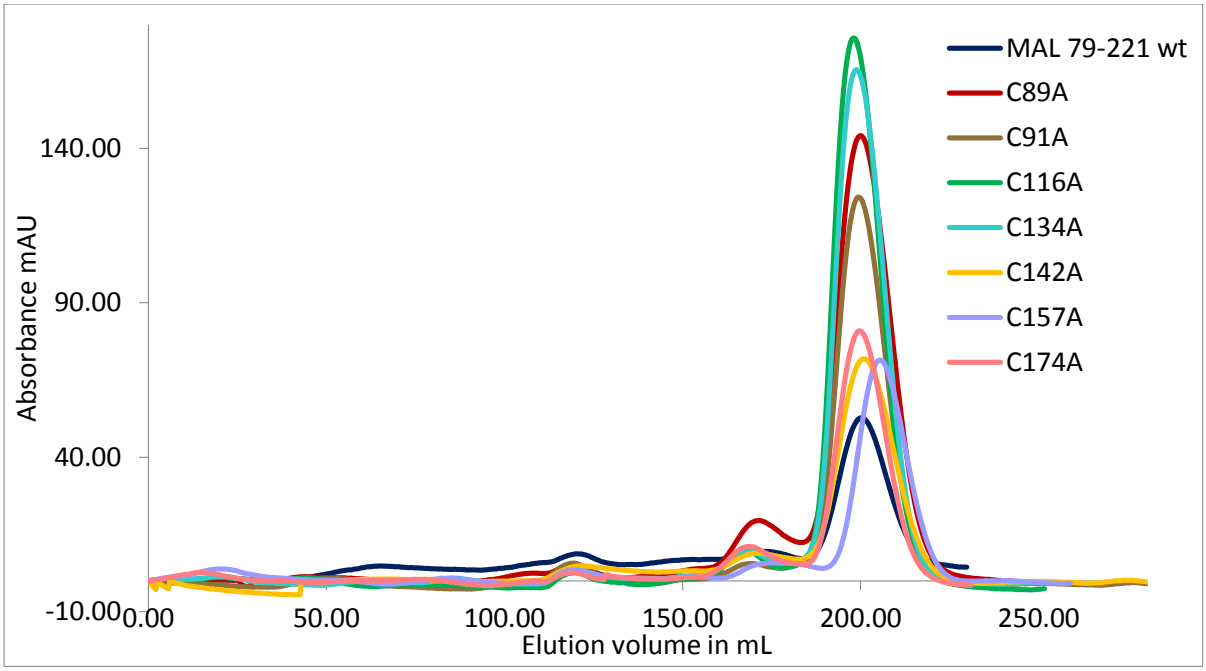


Figure 2.5 Size-exclusion chromatography curves of wild-type MAL^{TIR} and cysteine mutants following expression in one litre of *E. coli* BL-21 cells on a gel filtration column [Superdex 75 26/600, GE]. All mutants showed a two-peak elution profile with peaks at around 170 and 200 mL.

2.2.4 *Purification of MAL^{TIR} single-cysteine mutants expressed in E. coli SHuffle cells*

The size-exclusion chromatograms, using a smaller volume analytical column [Superdex 75 10/300, GE], of the MAL^{TIR} single cysteine-to-alanine mutants showed distinct differences between the mutants following expression in SHuffle cells (Figure 2.6). Protein expression using one litre of LB media revealed that mutants C89A and C116A produced relatively high yields compared with the rest of the mutants. Mutants C89A, C91A, C116A, C142A and C174A produced a two peak profile with a large peak eluting at around 13 mL and the smaller peak eluting at around 11 mL, which was consistent with the wild-type protein. Similarly, C157A also showed a two-peak profile, but contained a shifted elution peak. Comparatively, C134 showed a shoulder peak at 11 mL, however did not elute a major protein peak at 13 mL as the other mutants. Interestingly, expression using the SHuffle cell type consistently produced larger shoulder peak from around 7 mL compared to BL-21 cells.

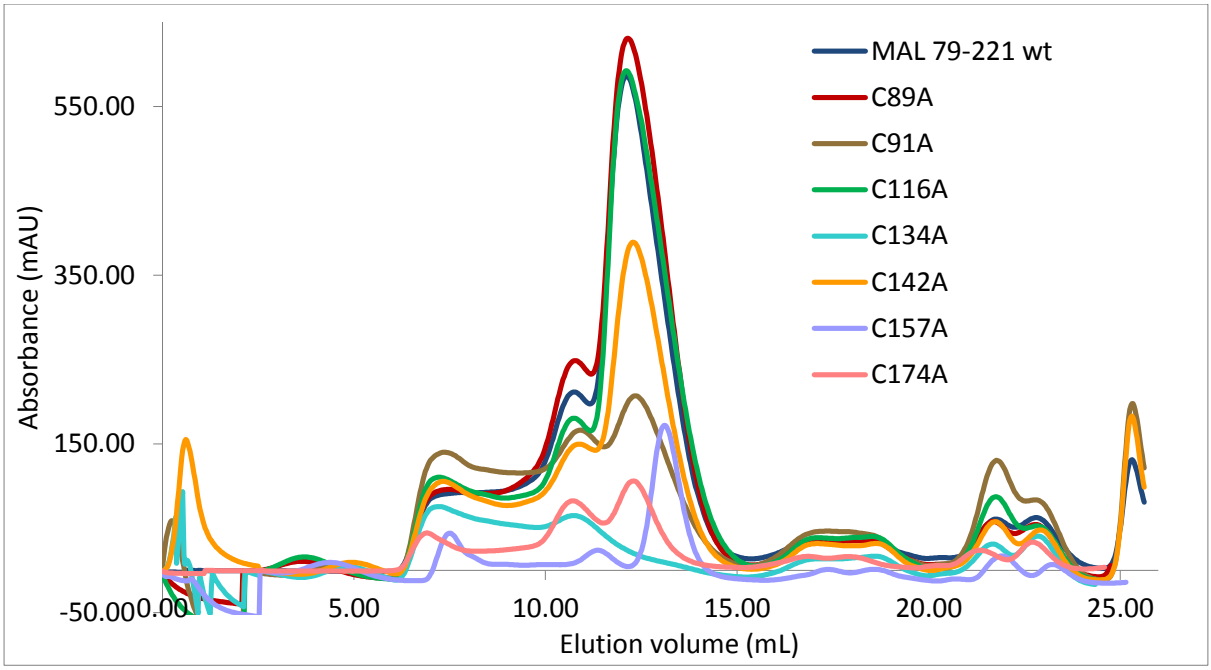


Figure 2.6 Size-exclusion chromatography curves of wild-type MAL^{TIR} and cysteine mutants following expression in one litre of *E. coli* SHuffle cells on the analytical gel-filtration column [Superdex 75 10/300, GE]. Cysteine-to-alanine mutants show varying elution and retention times between mutants.

2.2.5 *The redox state of the MAL^{TIR} cysteine residues varies in cysteine-to-alanine mutants following expression in E. coli SHuffle cells*

MAL^{TIR} mutants C89A, C116A and C174A all showed 3 oxidised and 3 reduced cysteine residues when expressed in *E. coli* SHuffle cells (Table 2.2). Interestingly, mutating cysteine position 91 resulted in 5 oxidised and 1 reduced cysteine, while mutating C142 favoured reduction with only 2 oxidised and 4 reduced cysteine residues. C157A resulted in 41% of reduced cysteine residues in the protein, equating to 2.5 out of 6 reduced residues. This showed that some cysteine residues were more prone to oxidation than others. The Ellman's assay provides the average number of cysteine residues that have been oxidised as a percentage of all cysteines present. It is therefore possible that a proportion of a particular residue is oxidised while the remaining is reduced. To address this issue, mass spectrometry was utilised to better understand the residues that were modified and the proportion of its population that was undergoing redox change.

Table 2.2 Number of reduced and oxidised cysteine residues in the MAL^{TIR} mutants following production in *E. coli* SHuffle cells.

MAL^{TIR} protein type	Oxidised cysteines	Reduced cysteines
wild-type	3	4
C89A	3	3
C91A	5	1
C116A	3	3
C134A	-	-
C142A	2	4
C157A	3 – 4	2 – 3
C174A	3	3

2.2.6 *Mass spectrometry detects a disulfide bond in MAL^{TIR}*

Intact mass of purified wild-type MAL^{TIR} was analysed using TripleTOF 5600 [Sciex] and Orbitrap Elite [Thermo] mass spectrometers to determine the presence of disulfide bonds by comparing the measured mass of the protein to the theoretically calculated mass. The theoretical mass of reduced MAL^{TIR} was calculated to be 16006 Da and 16017 Da, for monoisotopic and average masses, respectively. The TripleTOF mass analyser detected two species with average masses of 16017 Da and 16015 Da, which correspond to the fully reduced MAL^{TIR}, as well as a MAL^{TIR} species containing one disulfide bond. Similarly, the Orbitrap Elite detected a monoisotopic mass of 16004 Da, which corresponds to the monoisotopic mass of MAL^{TIR} with one disulfide bond. This indicates that following expression in the reducing environment of *E.coli* BL-21 cells, purified MAL^{TIR} is present as a mixture of two species consisting of a fully reduced form and a species with one disulfide bond.

2.2.7 *Cysteine residues of the MAL^{TIR} domain are susceptible to modification in an oxidative environment*

Purified wild-type MAL^{TIR} from *E. coli* BL-21 cells was exposed to 5 mM oxidised glutathione (GSSG) and then subjected to a trypsin digest. The peptides were analysed using mass spectrometry to identify which cysteine residues could undergo modifications in a highly oxidised environment. Similarly to the Ellman's reagent, glutathione reacts with the reduced thiols of cysteine residues to form a disulfide bond. Unlike Ellman's reagent, which calculates the average number of oxidised

cysteine residues from the population by relying on fluorescence, mass spectrometry is able to detect both modified and unmodified population sub-types as a proportion of one another. Modification by glutathione was tested at pH 8.6, in which the protein was purified, as well as a physiologically more relevant pH of 7.5. A number of modifications were detected, including *S*-glutathionylation, dioxidation and dehydroalanine formation, shown in Figure 2.7 B–D. Studies from Wang *et al* (27), and also by Maozziconacci *et al* (28), suggest that the formation of dehydroalanine is readily formed when a cysteine residue is exposed to heat of 30 °C in the pH range of 7.0 to 9.0 (27, 28). In the experimental setup used in this thesis, it is highly likely that the exposure to this temperature at the tested pHs of 7.5 and 8.6 resulted in dehydroalanine formation through this mechanism. Taken together, these results show that the cysteine residues of MAL^{TIR} are susceptible to modifications in solution. NMR was utilised to closely investigate the redox conditions that are required for these modifications to take place (Section 3.3.6). Before proceeding with NMR studies, protein stability assays were carried out to find experimental conditions that would provide NMR spectra of suitable quality.

Table 2.3 Modifications to cysteine residues from wild-type MAL^{TIR} at a pH of 7.5 or 8.6 detected using mass spectrometry, following the addition of oxidised glutathione.

Cysteine residue	Tris buffer pH 8.6	HEPES buffer pH 7.5
89	Dehydroalanine	Dioxidation
91	Glutathionylation, Dioxidation	Glutathionylation, Dehydroalanine
116	Glutathionylation	Glutathionylation
134	Dehydroalanine	Glutathionylation, Dehydroalanine
142	Glutathionylation, Dioxidation	Glutathionylation, Dioxidation
157	Glutathionylation	Glutathionylation
174	Glutathionylation	Glutathionylation

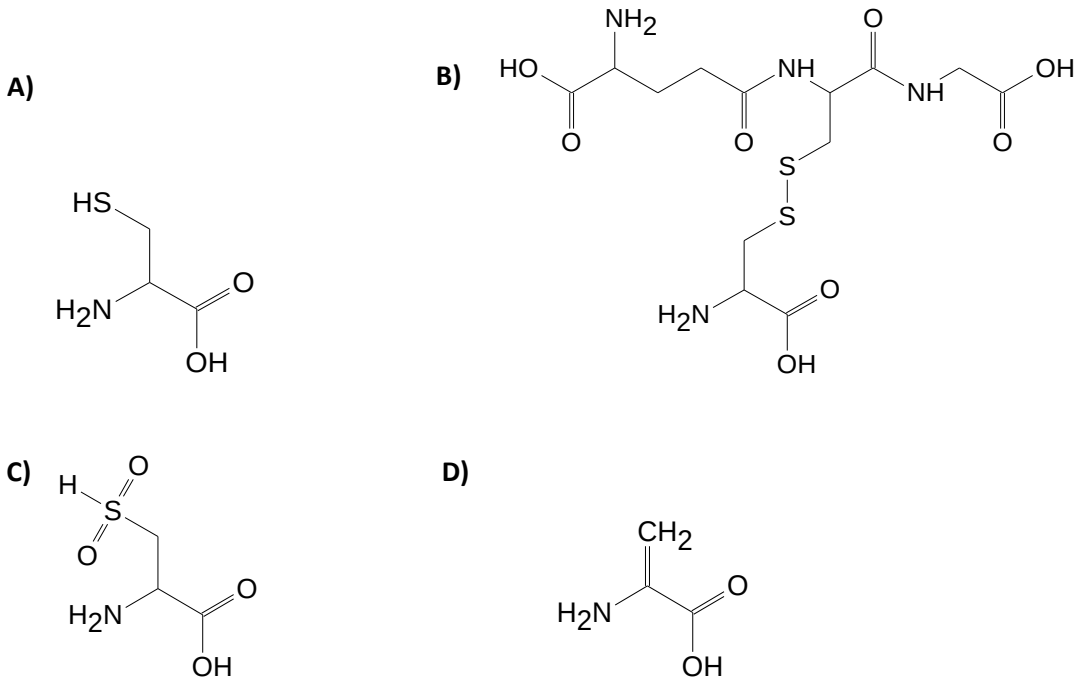


Figure 2.7 Structure of the amino-acid cysteine (**A**), and the modifications, *S*-glutathionylation (**B**), dioxidaton (**C**) and conversion to dehydroalanine (**D**).

2.2.8 *Protein solubility is dependent on temperature and concentration*

Solubility assays using the wild-type MAL^{TIR} were performed to find an appropriate condition for NMR experiments. The variables tested included temperature and concentration. The temperatures 4, 10, 18, 20, 25 and 37°C were tested, while the concentration ranged from 1.0 to 10.5 mg/mL. Overall, the data indicated that the protein was less prone to precipitation at lower concentrations than at higher concentrations (Figure 2.8). Similarly, the protein was less prone to precipitation at lower temperatures than at higher temperatures (Figure 2.9). Below 10°C, there was no precipitation at concentrations below 5.0 mg/mL. By contrast, at a temperature of 25°C, the precipitation rate increased with rising protein concentration.

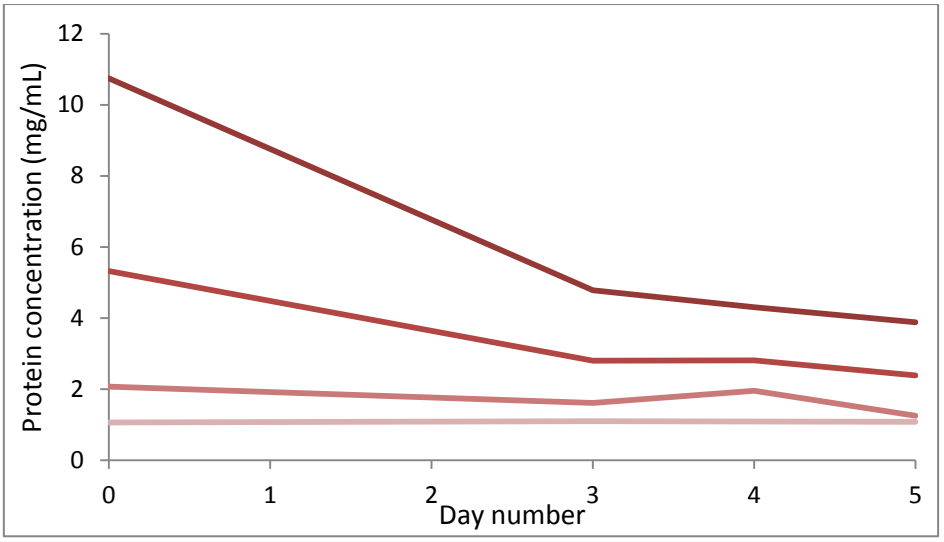


Figure 2.8 Wild-type MAL^{TIR} concentrations at 25°C over a period of 5 days. The graph indicates that a higher rate of protein loss is observed at increased concentrations.

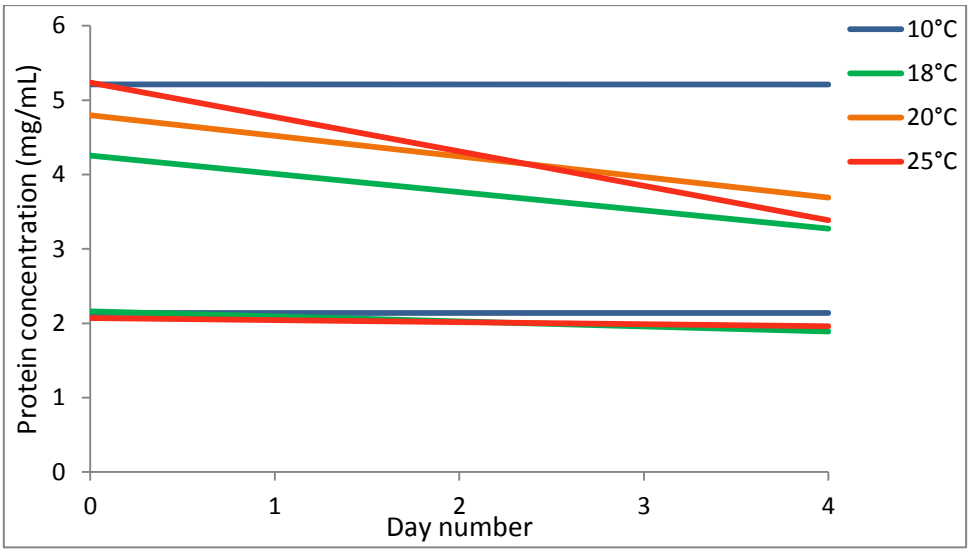


Figure 2.9 Wild-type MAL^{TIR} concentration at varying temperatures. The graph indicates that protein concentration is more stable at lower temperatures.

2.2.9 *Determining a suitable condition for NMR spectroscopy using ^{15}N-labelled* wild-type *MAL^{TIR}*

A compromise between temperature and protein concentration was required to maintain high signal intensity for NMR experiments, while keeping the protein stable in solution. Protein expression was carried out as described in Section 2.2.1 using *E. coli* BL-21 cells, however in order to introduce isotope labels, the cells were grown in minimal (M9) media (0.1 M KH_2PO_4, 0.45 M Na_2HPH_4, 0.09 M NaCl, 1× Vitamin solution [ThermoFisher], 3.75×10^{-6} M Thiamine, 8×10^{-4} M $MgSO_4$, 4×10^{-5} M $CaCl_2$, 9×10^{-3} M $^{15}NH_4Cl$, 0.01 M D-glucose), which contained 100 mg/L of ampicillin as well as ^{15}N-labelled ammonium chloride as the sole source of nitrogen. Protein expression was induced using 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and the temperature was reduced to 20°C overnight. The protein was purified as described in Section 2.2.1, using nickel affinity and size-exclusion chromatography. Using the purified ^{15}N-labelled MAL^{TIR} protein, three conditions (5 mg/mL at 10°C, 2.5 mg/mL at 10°C and 2.5 mg/mL at 18°C) were selected based on protein solubility data from Section 2.2.8 and the properties of NMR, which indicate that at higher temperatures the molecular correlation time is shorter and therefore produces narrower line-widths leading to a more intense signal. On the other hand, at lower temperatures, the tumbling rate slows down to produce broader lines that are less intense; although at lower temperatures, the amide NH exchange rate with solution H^+ is slowed down, which improves signal intensity. The protein was measured using 5 mm susceptibility matched NMR cells (Shigemi Inc.) and NMR data were acquired using a Bruker Avance II spectrometer equipped with a cryogenically cooled probe and operating at a nominal ^{1}H frequency of 900 MHz. Data was visualised and analysed using CCPNMR software (29). NMR spectra showed that the signal improved when the concentration and temperature increased. Visually, the spectra between 5 mg/mL at 10°C and 2.5 mg/mL at 18°C were very similar, with the higher concentration showing slightly more signals, (Figure 2.10). The signal difference between 2.5 mg/mL at 10 and 18°C was notable. The drop in temperature by 8°C produced less signal, which was visually represented by less peaks in the spectrum. The concentration of 2.5 mg/mL at a temperature of 18°C was chosen as the best condition that would lay the foundation for future NMR work.

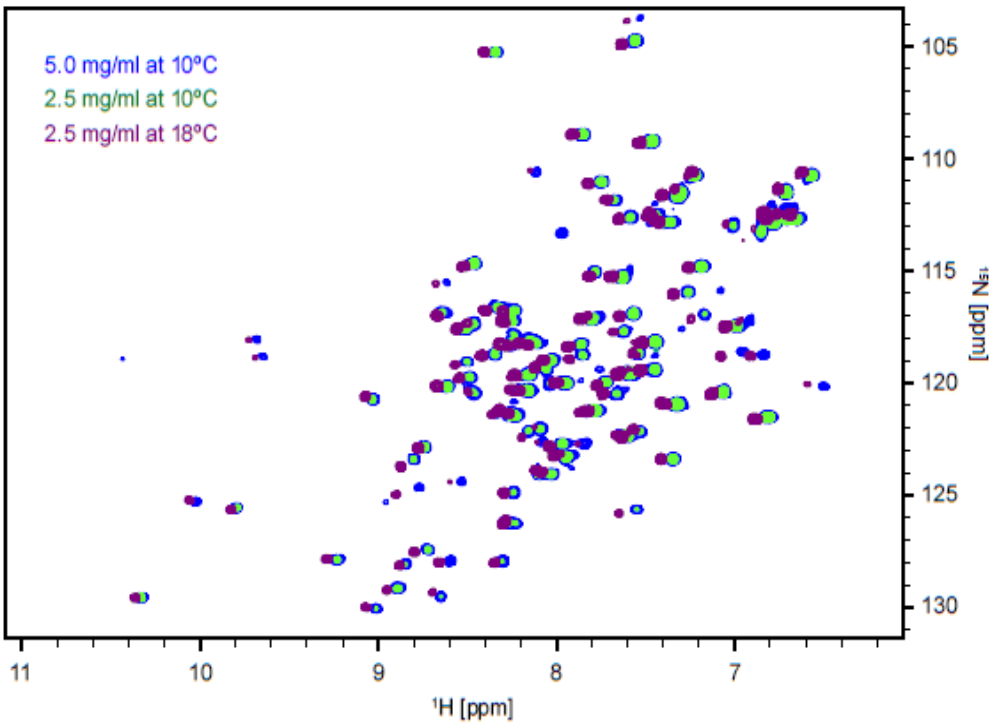


Figure 2.10 Overlay of ^{1}H-^{15}N-HSQC spectra of ^{15}N-labelled wild-type MAL^{TIR} at: 5 mg/mL at 10°C (blue), 2.5 mg/mL at 10°C (green) and 2.5 mg/mL at 18°C (purple).

2.2.10 *Determining the location of the cysteine residues in the ^{1}H-^{15}N-HSQC spectrum*

The wild-type MAL^{TIR} was expressed in *E. coli* BL-21 cells, which were grown in M9 media (Section 2.2.9) with the addition of ^{15}N-labelled ammonium chloride as the sole nitrogen source to produce ^{15}N-labelled MAL^{TIR}. In addition, the protein was also expressed in cysteine auxotroph *E. coli* bacteria where the media was supplemented with unlabelled L-cysteine, as these cells are unable to synthesise this amino acid. This resulted in a ^{15}N-labelled amide-backbone of the protein, with unlabelled cysteine amino-acids. An overlay of two ^{1}H-^{15}N-HSQC spectra obtained from protein expression from these cell types revealed the location of the cysteine residues (Figure 2.11). The protein that was expressed in *E. coli* BL-21 cells (red) revealed peaks that represent ^{1}H-^{15}N correlation from atoms corresponding to the amide backbone and side-chains of the MAL^{TIR} protein. By contrast, the protein that was purified from the *E. coli* cysteine auxotroph strain and supplemented with unlabelled cysteine showed six less resonances (blue). One cysteine residue could not be identified.

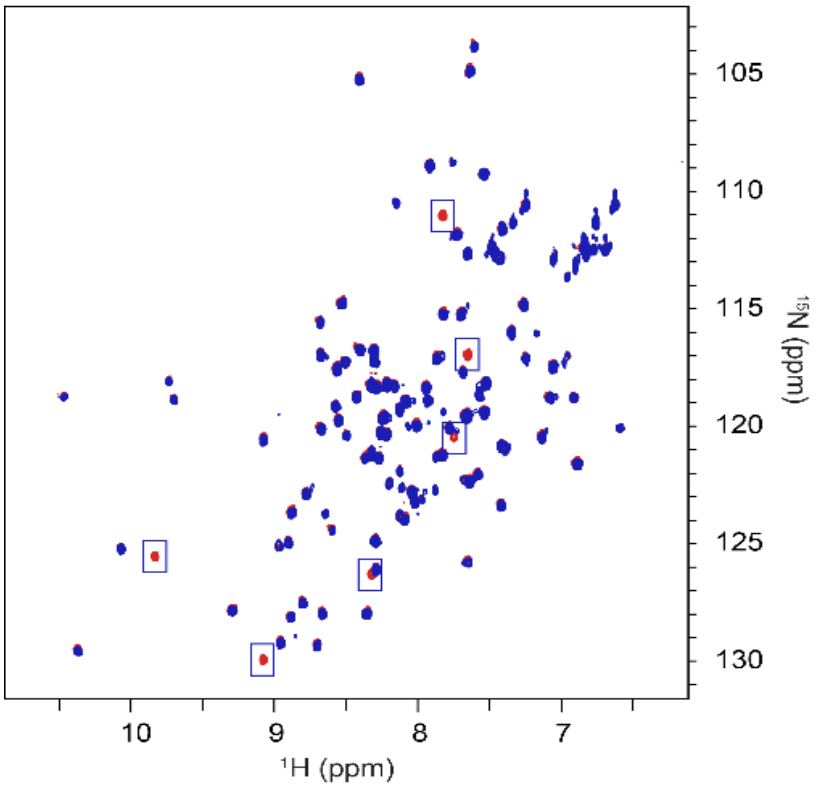


Figure 2.11 Overlay of ^{1}H-^{15}N-HSQC spectra of ^{15}N-labelled wild-type MAL^{TIR} with labelled (red) and unlabelled cysteine amino acids (blue). The positions of six cysteine amino acids are indicated by boxes.

2.2.11 *MAL^{TIR} precipitation is reversible*

Due to the high sensitivity of the ^{1}H-^{15}N-HSQC experiment, a low concentration of 2.5 mg/mL was used; however, at this concentration, the less sensitive triple resonance experiments produced insufficient signal intensities and the concentration was increased to 5 mg/mL. From the earlier protein stability assays, it was clear that the protein would precipitate more rapidly at higher temperatures at increasing concentrations. In light of this, the protein was kept on ice prior to and between experiments, which resulted in the recovery of precipitated protein. A schematic of this reversible process is shown in Figure 2.12 A–E.

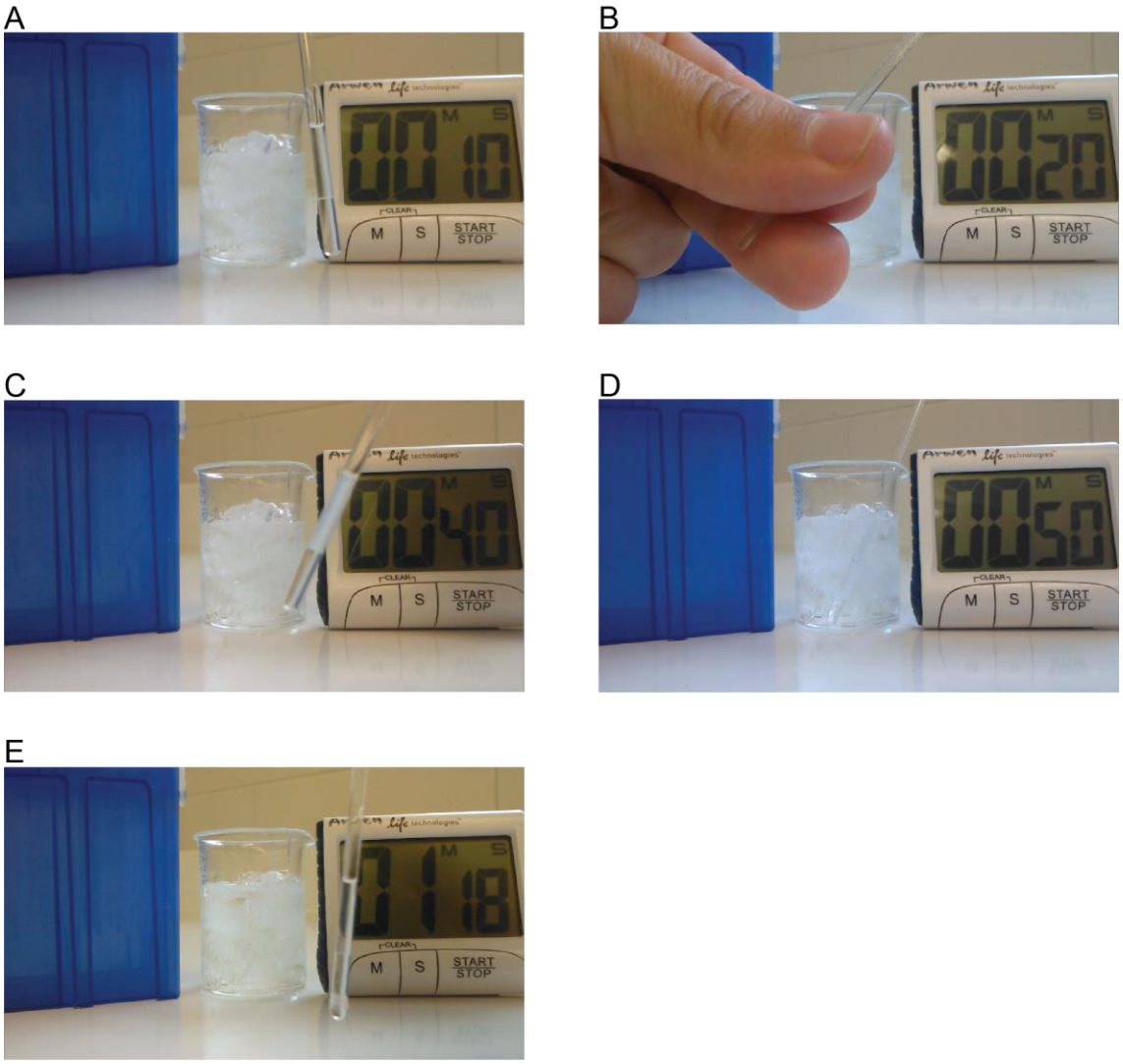


Figure 2.12 Visual representation showing that the precipitation of MAL^{TIR} following exposure to heat is reversible. Image A shows soluble protein at room temperature, which is then exposed to ~37°C for 10 seconds in image B. This results in the precipitation of the protein as shown in image C. Following exposure to 0°C in image D, the protein once again re-solubilises (image E).

2.2.12 *Chemical shifts of the cysteine residues*

The wild-type MAL^{TIR} amide backbone atoms (^{15}N, $^{13}C_\alpha$, $^{13}C_\beta$, $^{13}C'$, H_N, H_α, H_β) were assigned using triple resonance three-dimensional NMR spectra, which included, HNCACB, CBCA(CO)NH, and HNCO. Given that protein precipitation was reversible, the experiments were conducted at 18°C at 5 mg/mL overnight and placed in the 4°C fridge during the day to re-solubilise. Sample recovery was confirmed by ^{1}H-^{15}N-HSQC experiments and overlayed with the initial ^{1}H-^{15}N-HSQC to show that the protein was unchanged. Although this process of reviving the precipitated protein was used, wild-type MAL^{TIR} still provided only limited data due to its solubility over a limited time span of 12 hours.

Using the CBCA(CO)NH experiment, the chemical shifts of the carbon-β atoms on the cysteine residues were determined. Early work by Sharma *et al* in 2000 (30), and later validated by Martin *et al* in 2010 (31), showed that the chemical shift of the carbon-β atom is highly sensitive to the redox state of the neighbouring sulfur atom, and could therefore be used to determine the residue's redox state (30). A carbon-β shift greater than 34.5 ppm correlated with oxidised state, whereas a shift less than 32.5 ppm was consistent with a reduced cysteine residue (30). Chemical shifts between 32.5 and 34.5 ppm could be reduced or oxidised, depending on the secondary structure location of the cysteine and neighbouring residues. Six of the cysteines were identified on the ^{1}H-^{15}N-HSQC (Section 2.2.10), while the seventh cysteine was identified at a lower pH of 7.5 following experiment in Section 3.3.5, which revealed extra peaks. In the wild-type MAL^{TIR} protein, six of the seven cysteine residues contained a chemical shift that was consistent with the reduced redox state, shown in Table 2.4, while one cysteine (C116) was undefined with a carbon-β shift of 33.4 ppm. This result is consistent with mass spectrometry data (Section 2.2.6), which detected a mass corresponding to a fully reduced species and another species containing one disulfide bond. Part one of this chapter has indicated that the cysteines of MAL^{TIR} are susceptible to change following different redox conditions. Further structural analysis and physiological relevance is discussed in Chapter 3.

Table 2.4 Cysteine carbon-β shifts of wild-type MAL^{TIR} following expression in BL-21 *E. coli* cells

Cysteine residue	Carbon β shift (ppm)	Redox state (<32.5 ppm = reduced)
89	28.3	Reduced
91	28.5	Reduced
116	33.4	Reduced or oxidised
134	29.5	Reduced
142	31.2	Reduced
157	28.5	Reduced
174	29.0	Reduced

2.3 PART 2: Dimer formation of MAL^{TIR}

The purification of the wild-type MAL^{TIR} domain from *E. coli* BL-21, Shuffle and Origami cells all displayed a similar size-exclusion elution chromatogram, which depicted a higher molecular weight product eluting earlier at 170 mL, as compared to the major product, which eluted at 200 mL (Section 2.2.1). Repurifying this first peak at 170 mL over size-exclusion chromatography elutes a single peak at a volume of 170 mL. This suggests that the MAL^{TIR} protein could be forming a stable higher oligomeric state.

2.3.1 *MAL^{TIR} oligomer is not covalently bonded*

Following size-exclusion chromatography of MAL^{TIR} (Section 2.2.1), a higher molecular weight (HMW) peak, which elutes at 170 mL, was run on 12% SDS PAGE with or without reducing agent (DTT) and compared to the major elution peak at 200 mL that was confirmed to be the monomer of MAL^{TIR} (Figure 2.13). All lanes showed a major band at an estimated 15 kDa, which corresponded to the expected size for the monomeric MAL^{TIR} protein. A faint band is observed at around 25 kDa in samples corresponding to the smaller elution peak 170 mL, suggesting a stable dimeric product. Overall, no significant changes occurred with and without the addition of reducing agent, suggesting that dimer or oligomer formation between MAL^{TIR} monomers is not likely to be stabilised by covalent bonds between cysteine residues.

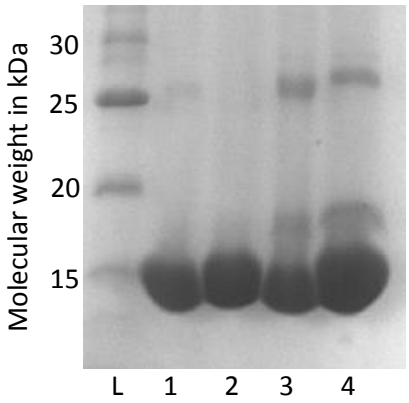


Figure 2.13 SDS PAGE of wild-type MAL^{TIR} monomer and higher molecular weight (HMW) species, reduced and non-reduced. **1)** Monomer, **2)** Monomer with DTT, **3)** HMW species, **4)** HMW species with DTT.

2.3.2 *Higher molecular weight species of the MAL^{TIR} domain is a salt independent interaction*

To investigate the effect of salt on the ability of MAL^{TIR} to form stable homo-dimers, the protein was separated by size-exclusion chromatography using the analytical gel-filtration column [GE] in either no salt or a salt concentration of 500 mM. The salt concentration had no significant effect on the elution profile of MAL^{TIR}, with both chromatograms showing a very similar elution profile (Figure 2.14).

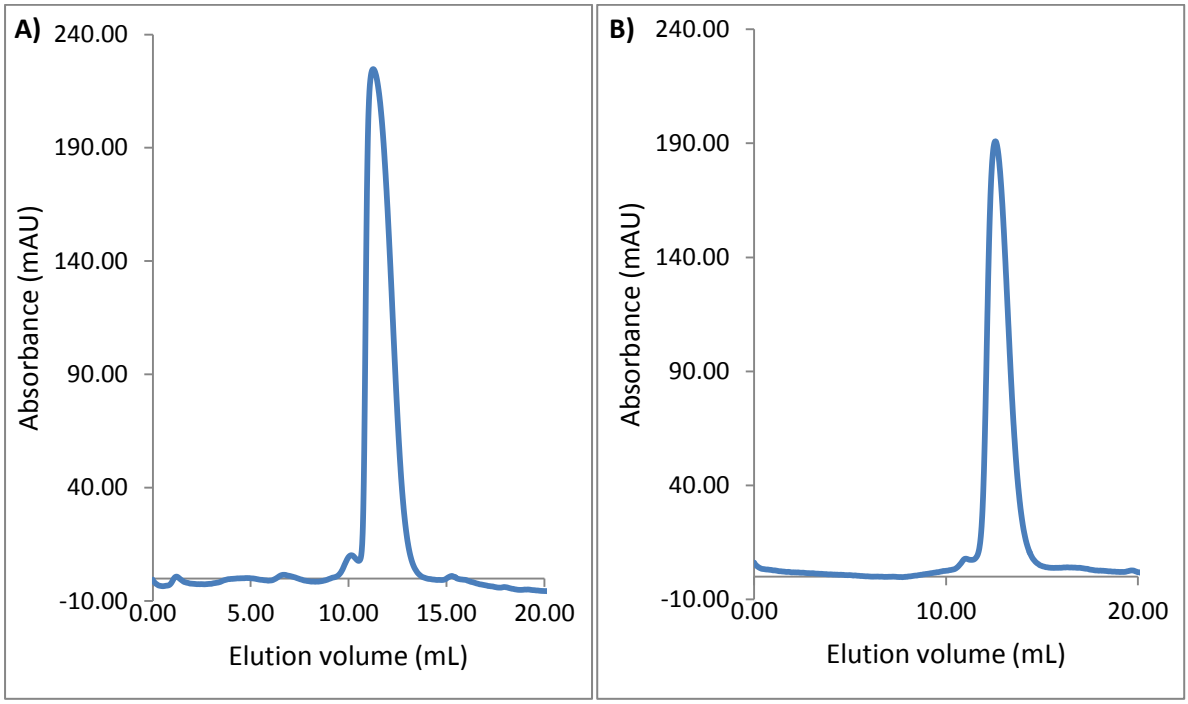


Figure 2.14 Size-exclusion chromatogram of MAL^{TIR} purified with either: **A)** 0 mM salt or **B**) 500 mM salt.

2.3.3 *Size and shape analysis the MAL^{TIR} domain monomer and higher molecular weight moiety*

Both smaller shoulder peak and the major elution peaks from size-exclusion chromatography were analysed by small angle X-ray scattering (SAXS) to confirm the size of the higher molecular weight elution product. The molecular weight calculated from the scattering of the major peak was consistent with the monomer, and was compared to the crystal structure determined by Valkov *et al* (15) using FoXS (32) (Figure 2.15). The goodness-of-fit is evaluated by a χ-metric, which is the normalised sum of the squares of the residuals of the predicted and experimental data, divided by the experimental errors (32). A χ-value of 1 constitutes a perfect fit to the data, although this is difficult to achieve in practice due to challenges in estimation of the errors (33, 34). In this case, the MAL^{TIR} crystal structure alone is not sufficient to reproduce the scattering (χ = 3.8), although the

agreement was improved by the addition of the missing AB loop (χ = 2.4). However, the fits are close, suggesting that the predominant members of the solution ensemble are similar in shape.

Furthermore, the smaller shoulder peak that elutes before the monomer (Figure 2.3), corresponded to the molecular weight for a MAL^{TIR} dimer. The dimeric arrangement using two monomer interfaces in the crystal was not consistent with the scattering (χ = 13.54 and 15.30) (Figure 2.16). Rigid body modelling using two copies of the crystallographic monomer improved this to χ = 4.31, but an ideal fit was still not achieved (Figure 2.17). This indicates that the dimer interface of the purified dimer is different to that observed between the monomers in the protein crystal, but also that minor conformational changes in the monomer may be present.

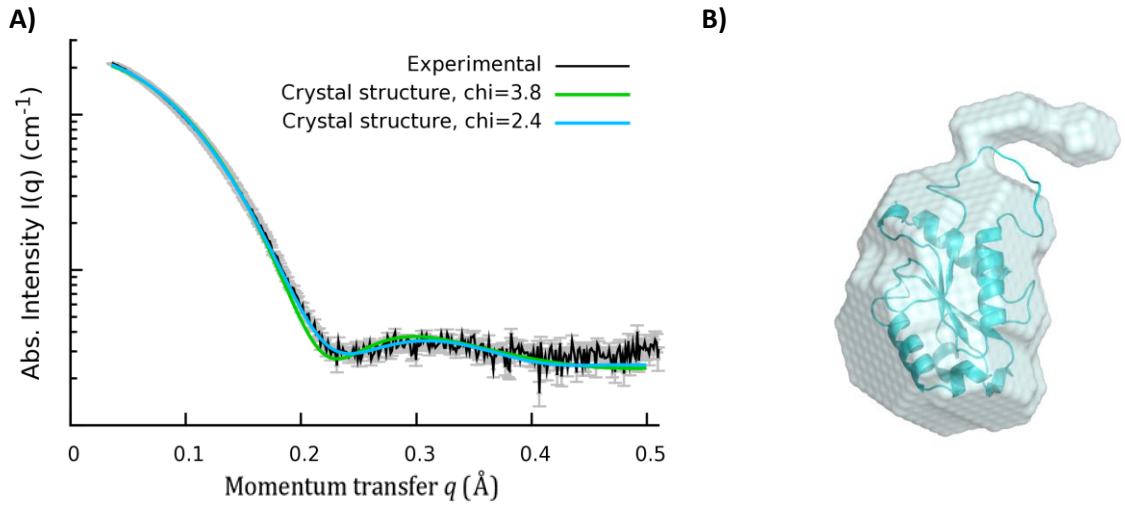


Figure 2.15 A) SAXS curves showing the experimental plot of MAL^{TIR} monomer with (blue) and without (green) added loop; chi= 2.4 and 3.8, respectively. **B)** Rigid body modelling of SAXS data using the MAL^{TIR} monomer crystal structure from Valkov *et al* (15).

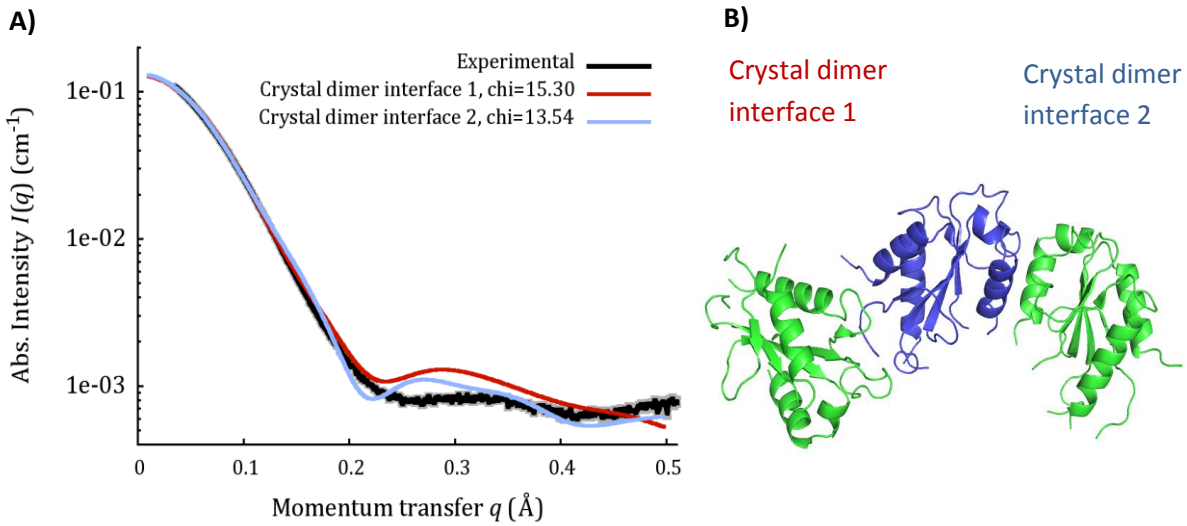


Figure 2.16 A) Comparison of the experimentally measured SAXS curve (black) and two theoretically generated curves based on two crystal interfaces from protein crystal of MAL^{TIR} (red and blue). **B)** Visual representation of the two dimer interfaces detected in the protein crystal of MAL^{TIR} (15).

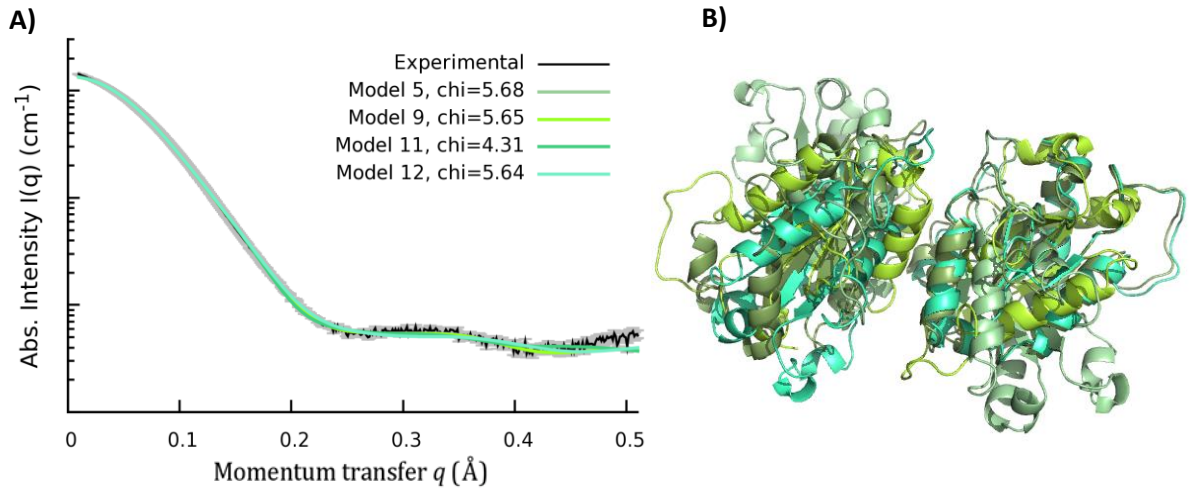


Figure 2.17 A) Experimental SAXS curve (black) of MAL^{TIR} dimer with different models for the interaction interface. **B)** Model of different dimer arrangements that fit experimental SAXS data. Colours represent different models.

2.3.4 *Crystallisation of the MAL^{TIR} dimer*

Evidence from SAXS confirmed that a dimeric species was present, however the nature of the dimer interface remained unclear. Crystallography was attempted to determine the structure of the MAL^{TIR} dimer. MAL^{TIRC116A} mutant was used for screening as this mutant produced a higher yield of this dimer compared to the wild-type (Section 2.2.3). Initial grid screen conditions were tested using commercially available screens including: PEG/Ion and PEGRx [Hampton Research] and Pact Premier and JCSG+ [Molecular Dimensions]. Grid screens were then designed around the most promising conditions. These promising conditions were tested using varying precipitant, pH, salt, and reducing agent (Table 2.5). One condition that was composed of 10% PEG 1000 with 1 mM DTT, produced small needle-like crystals following 10 days incubation at 18°C, which are shown in Figure 2.18. These crystals were too small to test diffraction by X-rays, and in conjunction with the small yield following purification of the dimer, made it difficult to study.

Table 2.5 Conditions tested for the crystal formation of the MAL^{TIR} dimer

Condition	Range
PEG type	400, 1000, 3500, 6000, 8000
PEG concentration	0 – 12 %
pH	6 – 10
NaCl	0 – 200 mM
$MgSO_4$	150 – 600 mM
Metal additives	$MgSO_4$, $MgCl_2$, $ZnCl_2$, $ZnSO_4$
Methods	Hanging drop, sitting drop, streak seeding
Protein to buffer drop ratio	3:1, 2:1, 1:1
DTT	With/ without

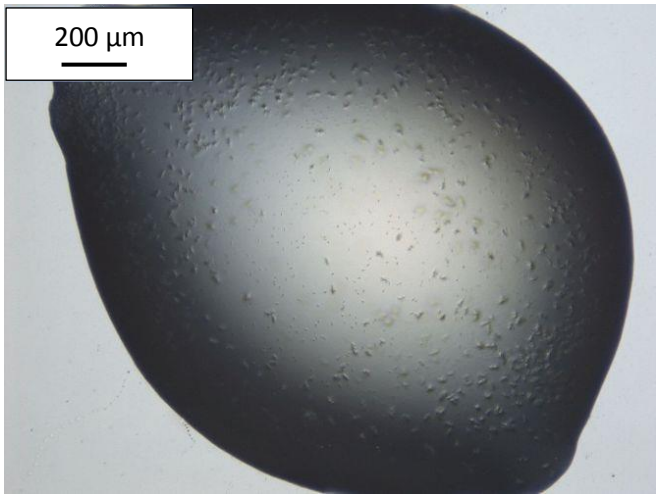


Figure 2.18 Small needle-like crystals of $MAL^{TIRC116A}$ dimer obtained in 10% polyethylene glycol 1000 with 1 mM DTT following 10 days incubation at 18°C.

2.4 Summary: Proposed model for MAL^{TIR} oligomerisation

Based on the compilation of results presented in this chapter, the human MAL^{TIR} exists in a dynamic equilibrium in solution between a monomer and higher oligomeric state (Figure 2.19). The formation of the oligomeric state is dependent on a combination of protein concentration (> 2 mg/mL) and temperature (> 10°C) at a neutral pH. These conditions are inter-related and the equilibrium can be shifted by changing any one of these conditions. Following the purification of soluble MAL^{TIR}, around 10% of protein purifies as an irreversible dimer. The monomeric interface of this dimer is unknown, however, it is unlikely that covalent interactions between cysteine thiols play a role. In addition, salt concentrations of zero and 500 mM do not appear to influence dissociation of the dimer.

Results presented here are consistent with the paradigm of higher order assembly formation (35-38). The formation of an oligomeric complex requires a seed and a favourable condition that stabilises further polymerisation. The formation of the oligomer is determined by a sigmoidal curve that is a function of concentration (35). The stable MAL^{TIR} dimer presented in this chapter could suggest a role for this molecule as the initial seed, upon which the oligomer can build when conditions are favourable to stimulate a response. For example, the study by Tay *et al* (39) has shown that activation is induced once a critical mass is reached, initiating an all-or-nothing response. Tay *et al* reported on tumor necrosis factor-alpha (TNF-α)-induced NF-κB activation using TNF-α concentrations that spanned over four orders of magnitude (39). NF-κB was activated once a critical threshold of TNF-α was reached, with a greater number of mammalian cells switching on a response at higher concentrations of TNF-α (39). Further to this, higher order assemblies amplify signal while

reducing biological noise. The higher order assembly paradigm suggests that without larger assemblies, the signal transduction is thermodynamically and kinetically less efficient with transient fluctuations in concentration, diffusion, and protein conformation (35).

A striking feature of the MAL^{TIR} oligomer demonstrated in this chapter (Figure 2.12) shows its ability to reverse into soluble monomers. This novel feature of MAL^{TIR} suggests a possible regulation of the signalling mechanism for the protein. Association of TIR domains has been reported to promote signalling, such as in TLRs. A study done by Ve *et al* (unpublished) confirmed that MAL formed the oligomer by fitting monomers into an electron density map produced by electron microscopy. A similar feature for reversible oligomers was first reported by Motshwene *et al* (40), between downstream protein MyD88 and its binding partners IRAKs; a complex that was later crystallized by Lin *et al* (41). This interaction constitutes death-domain interactions, which form a helical assembly at increasing concentrations. Interestingly, MyD88 is also prone to self-association. This process could be controlled with the addition of 50 mM salt and reducing agent (40). Given that the downstream signalling protein MyD88 reversibly self-associates and forms a complex with IRAK proteins 4 and 2, it is likely that the reversible oligomer observed in MAL^{TIR} is a physiologically relevant event, providing a scaffold for efficient downstream signalling.

The observations made here are critical, particularly early in TLR signalling during the TIR domain interactions of TLR, MAL, and MyD88. Higher-order assemblies between these proteins, especially MAL as shown through the results in this chapter, would allow for an efficient transfer and amplification of signalling to downstream molecules such as the IRAK complex.

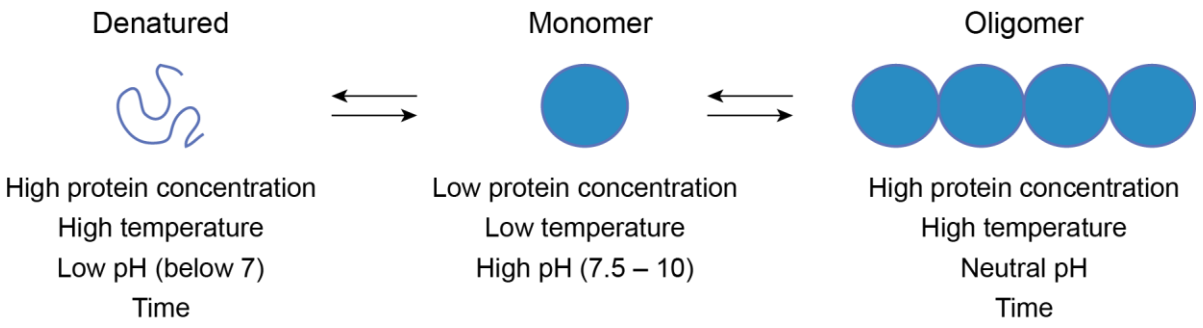


Figure 2.19 A model showing conditions that determine the dynamic equilibrium of the monomeric and dimeric states of the MAL^{TIR} protein, as well as conditions that lead to the irreversible denaturation of the protein.

2.5 Chapter 2 references

1. Takeda, K., and Akira, S. (2005) Toll-like receptors in innate immunity, *International Immunology* **17**, 1–14.
2. Medzhitov, R. (2001) Toll-like receptors and innate immunity, *Nature Reviews Immunology* **1**, 135–145.
3. Choi, Y. J., Jung, J., Chung, H. K., Im, E., and Rhee, S. H. (2013) PTEN regulates TLR5-induced intestinal inflammation by controlling Mal/TIRAP recruitment, *F ASEB Journal* **27**, 243–254.
4. Yamamoto, M., Sato, S., Hemmi, H., Sanjo, H., Uematsu, S., Kaisho, T., Hoshino, K., Takeuchi, O., Kobayashi, M., Fujita, T., Takeda, K., and Akira, S. (2002) Essential role for TIRAP in activation of the signalling cascade shared by TLR2 and TLR4, *Nature* **420**, 324–329.
5. Horng, T., Barton, G. M., Flavell, R. A., and Medzhitov, R. (2002) The adaptor molecule TIRAP provides signalling specificity for Toll-like receptors, *Nature* **420**, 329–333.
6. Horng, T., Barton, G. M., and Medzhitov, R. (2001) TIRAP: an adapter molecule in the Toll signaling pathway, *Nature Immunology* **2**, 835–841.
7. Fitzgerald, K. A., Palsson-McDermott, E. M., Bowie, A. G., Jefferies, C. A., Mansell, A. S., Brady, G., Brint, E., Dunne, A., Gray, P., Harte, M. T., McMurray, D., Smith, D. E., Sims, J. E., Bird, T. A., and O'Neill, L. A. (2001) Mal (MyD88-adapter-like) is required for Toll-like receptor-4 signal transduction, *Nature* **413**, 78–83.
8. Kagan, J. C., and Medzhitov, R. (2006) Phosphoinositide-mediated adaptor recruitment controls Toll-like receptor signaling, *Cell* **125**, 943–955.
9. Ve, T., Gay, N. J., Mansell, A., Kobe, B., and Kellie, S. (2012) Adaptors in toll-like receptor signaling and their potential as therapeutic targets, *Current Drug Targets* **13**, 1360–1374.
10. Nagpal, K., Plantinga, T. S., Wong, J., Monks, B. G., Gay, N. J., Netea, M. G., Fitzgerald, K. A., and Golenbock, D. T. (2009) A TIR domain variant of MyD88 adapter-like (Mal)/TIRAP results in loss of MyD88 binding and reduced TLR2/TLR4 signaling, The *Journal of Biological Chemistry* **284**, 25742–25748.
11. Ferwerda, B., Alonso, S., Banahan, K., McCall, M. B., Giamarellos-Bourboulis, E. J., Ramakers, B. P., Mouktaroudi, M., Fain, P. R., Izagirre, N., Syafruddin, D., Cristea, T., Mockenhaupt, F. P., Troye-Blomberg, M., Kumpf, O., Maiga, B., Dolo, A., Doumbo, O., Sundaresan, S., Bedu-Addo, G., van Crevel, R., Hamann, L., Oh, D. Y., Schumann, R. R., Joosten, L. A., de la Rua, C., Sauerwein, R., Drenth, J. P., Kullberg, B. J., van der Ven, A. J., Hill, A. V., Pickkers, P., van der Meer, J. W., O'Neill, L. A., and Netea, M. G. (2009) Functional and genetic evidence that the Mal/TIRAP allele variant 180L has been selected by providing protection against septic shock, *Proceedings of the National Academy of Sciences U S A* **106**, 10272–10277.

12. Ladhani, S. N., Davila, S., Hibberd, M. L., Heath, P. T., Ramsay, M. E., Slack, M. P., Pollard, A. J., and Booy, R. (2010) Association between single-nucleotide polymorphisms in Mal/TIRAP and interleukin-10 genes and susceptibility to invasive haemophilus influenzae serotype b infection in immunized children, *Clinical Infectious Diseases* **51**, 761–767.

13. Khor, C. C., Chapman, S. J., Vannberg, F. O., Dunne, A., Murphy, C., Ling, E. Y., Frodsham, A. J., Walley, A. J., Kyrieleis, O., Khan, A., Aucan, C., Segal, S., Moore, C. E., Knox, K., Campbell, S. J., Lienhardt, C., Scott, A., Aaby, P., Sow, O. Y., Grignani, R. T., Sillah, J., Sirugo, G., Peshu, N., Williams, T. N., Maitland, K., Davies, R. J., Kwiatkowski, D. P., Day, N. P., Yala, D., Crook, D. W., Marsh, K., Berkley, J. A., O'Neill, L. A., and Hill, A. V. (2007) A Mal functional variant is associated with protection against invasive pneumococcal disease, bacteremia, malaria and tuberculosis, *Nature Genetics* **39**, 523–528.

14. Castiblanco, J., Varela, D. C., Castano-Rodriguez, N., Rojas-Villarraga, A., Hincapie, M. E., and Anaya, J. M. (2008) TIRAP (MAL) S180L polymorphism is a common protective factor against developing tuberculosis and systemic lupus erythematosus, *Infection, Genetics and Evolution* **8**, 541–544.

15. Valkov, E., Stamp, A., DiMaio, F., Baker, D., Verstak, B., Roversi, P., Kellie, S., Sweet, M. J., Mansell, A., Gay, N. J., Martin, J. L., and Kobe, B. (2011) Crystal structure of Toll-like receptor adaptor MAL/TIRAP reveals the molecular basis for signal transduction and disease protection, *Proceedings of the National Academy of Sciences of the U S A* **108**, 14879–14884.

16. Woo, J. R., Kim, S., Shoelson, S. E., and Park, S. (2012) X-ray Crystallographic Structure of TIR-Domain from the Human TIR-Domain Containing Adaptor Protein/MyD88-Adaptor-Like Protein (TIRAP/MAL), *Bulletin of the Korean Chemical Society* **33**, 3091–3094.

17. Snyder, G. A., Deredge, D., Waldhuber, A., Fresquez, T., Wilkins, D. Z., Smith, P. T., Durr, S., Cirl, C., Jiang, J., Jennings, W., Luchetti, T., Snyder, N., Sundberg, E. J., Wintrode, P., Miethke, T., and Xiao, T. S. (2014) Crystal structures of the Toll/Interleukin-1 receptor (TIR) domains from the Brucella protein TcpB and host adaptor TIRAP reveal mechanisms of molecular mimicry, The *Journal of Biological Chemistry* **289**, 669–679.

18. Lin, Z., Lu, J., Zhou, W., and Shen, Y. (2012) Structural insights into TIR domain specificity of the bridging adaptor Mal in TLR4 signaling, *PLoS One 7*, e34202.

19. Menon, D., Coll, R., O'Neill, L. A. J., and Board, P. G. (2015) GSTO1-1 modulates metabolism in macrophages activated through the LPS and TLR4 pathway, *Journal of Cell Science* **128**, 1982–1990.

20. Li, Q., Harraz, M. M., Zhou, W., Zhang, L. N., Ding, W., Zhang, Y., Eggleston, T., Yeaman, C., Banfi, B., and Engelhardt, J. F. (2006) Nox2 and Rac1 regulate H(2)O(2)-dependent

recruitment of TRAF6 to endosomal interleukin-1 receptor complexes, *Molecular and Cellular Biology* **26**, 140–154.

21. Cruz, C. M., Rinna, A., Forman, H. J., Ventura, A. L., Persechini, P. M., and Ojcius, D. M. (2007) ATP activates a reactive oxygen species-dependent oxidative stress response and secretion of proinflammatory cytokines in macrophages, The *Journal of Biological Chemistry* **282**, 2871–2879.
22. Ve, T., Williams, S. J., and Kobe, B. (2015) Structure and function of Toll/interleukin-1 receptor/resistance protein (TIR) domains, *Apoptosis* **20**, 250–261.
23. Eschenfeldt, W. H., Lucy, S., Millard, C. S., Joachimiak, A., and Mark, I. D. (2009) A family of LIC vectors for high-throughput cloning and purification of proteins, *Methods in Molecular Biology* **498**, 105–115.
24. Lobstein, J., Emrich, C. A., Jeans, C., Faulkner, M., Riggs, P., and Berkmen, M. (2012) SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm, *Microbial Cell Factories* **11**, 56.
25. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures, *Protein Expression and Purification* **41**, 207–234.
26. Ellman, G. L., Courtney, K. D., Andres, V., Jr., and Feather-Stone, R. M. (1961) A new and rapid colorimetric determination of acetylcholinesterase activity, *Biochemical Pharmacology* **7**, 88–95.
27. Wang, Z., Rejtar, T., Zhou, Z. S., and Karger, B. L. (2010) Desulfurization of cysteine-containing peptides resulting from sample preparation for protein characterization by mass spectrometry, *Rapid Communications in Mass Spectrometry* **24**, 267–275.
28. Mozziconacci, O., Kerwin, B. A., and Schoneich, C. (2011) Reversible hydrogen transfer reactions of cysteine thiyl radicals in peptides: the conversion of cysteine into dehydroalanine and alanine, and of alanine into dehydroalanine, The *Journal of Physical Chemistry B* **115**, 12287–12305.
29. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline, *Proteins* **59**, 687–696.
30. Sharma, D., and Rajarathnam, K. (2000) 13C NMR chemcial shifts can predict disulfide bond formation, *Journal of Biomolecular NMR* **18**, 165–171.
31. Martin, O. A., Villegas, M. E., Vila, J. A., and Scheraga, H. A. (2010) Analysis of $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts of cysteine and cystine residues in proteins: a quantum chemical approach, *Journal of Biomolecular NMR* **46**, 217–225.

32. Schneidman-Duhovny, D., Hammel, M., and Sali, A. (2010) FoXS: a web server for rapid computation and fitting of SAXS profiles, *Nucleic Acids Research* **38**, W540–544.

33. Jacques, D. A., Guss, J. M., Svergun, D. I., and Trewhella, J. (2012) Publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution, *Acta Crystallogrphica Section D Biological Crystallography* **68**, 620–626.

34. Pelikan, M., Hura, G. L., and Hammel, M. (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering, *General Physiology and Biophysics* **28**, 174–189.

35. Wu, H. (2013) Higher-order assemblies in a new paradigm of signal transduction, *Cell* **153**, 287–292.

36. Park, H. H., Logette, E., Raunser, S., Cuenin, S., Walz, T., Tschopp, J., and Wu, H. (2007) Death domain assembly mechanism revealed by crystal structure of the oligomeric PIDDosome core complex, *Cell* **128**, 533–546.

37. Li, J., McQuade, T., Siemer, A. B., Napetschnig, J., Moriwaki, K., Hsiao, Y. S., Damko, E., Moquin, D., Walz, T., McDermott, A., Chan, F. K., and Wu, H. (2012) The RIP1/RIP3 necrosome forms a functional amyloid signaling complex required for programmed necrosis, *Cell* **150**, 339–350.

38. Yin, Q., Lin, S. C., Lamothe, B., Lu, M., Lo, Y. C., Hura, G., Zheng, L., Rich, R. L., Campos, A. D., Myszka, D. G., Lenardo, M. J., Darnay, B. G., and Wu, H. (2009) E2 interaction and dimerization in the crystal structure of TRAF6, *Nature Structural & Molecular Biology* **16**, 658–666.

39. Tay, S., Hughey, J. J., Lee, T. K., Lipniacki, T., Quake, S. R., and Covert, M. W. (2010) Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing, *Nature* **466**, 267–271.

40. Motshwene, P. G., Moncrieffe, M. C., Grossmann, J. G., Kao, C., Ayaluru, M., Sandercock, A. M., Robinson, C. V., Latz, E., and Gay, N. J. (2009) An oligomeric signaling platform formed by the Toll-like receptor signal transducers MyD88 and IRAK-4, *The Journal of Biological Chemistry 284*, 25404-25411.

41. Lin, S. C., Lo, Y. C., and Wu, H. (2010) Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling, *Nature* **465**, 885–890.

Chapter 3

Solution structure of the TLR adaptor MAL/TIRAP TIR domain

3.1 Contribution

Chapter 3 includes a manuscript titled, "*Solution structure of the TLR adaptor MAL/TIRAP TIR domain in the reduced state reveals large structural rearrangements upon disulfide-bond formation*," which will be submitted shortly for publication. In this piece of work, I expressed and purified all MAL^{TIR} constructs, prepared all samples for NMR spectroscopy, assigned ^{1}H, ^{15}N, ^{13}C resonances to wild-type MAL^{TIR} and the C116A mutant, refined the structure of $MAL^{TIRC116A}$, wrote the manuscript, and created the figures. All authors contributed to critical analysis and review of the manuscript.

Solution structure of the TLR adaptor MAL/TIRAP TIR domain in the reduced state reveals large structural rearrangements upon disulfide-bond formation

Peter Lavrencic[1,2], Mehdi Mobli[2], Bostjan Kobe[1]

[1]School of Chemistry and Molecular Biosciences, Institute for Molecular Bioscience and Australian Infectious Diseases Research Centre, University of Queensland, Brisbane, Queensland 4072, Australia. [2]Centre for Advanced Imaging, University of Queensland, Brisbane, Queensland 4072, Australia.

3.2 Introduction

Toll-like receptors (TLRs) serve as innate immune sensors for microbial stimuli such as bacterial lipids and nucleic acids (1). These microbial molecules bind to the extracellular domains of the receptors, inducing their oligomerization, including the association of their intramolecular TIR (Toll/interlekin-1 receptor/resistance protein) domains (2). The associated receptor TIR domains serve as a scaffold for the binding of TIR domain-containing adaptor proteins and effector enzymes in the cytosol, eventually leading to a broad range of inflammatory and antimicrobial responses, as well as the maturation of the adaptive immune response.

The ten human TLRs each use a specific combination of four TIR domain-containing adaptors to activate specific sets of signaling pathways (3). All TLRs except TLR3 activate the 'MyD88-dependent pathway', through binding the adaptor MyD88 (myeloid differentiation primary response gene 88), either directly or with the help of the bridging adaptor MAL (MyD88-adaptor-like), also known as TIRAP (TIR-domain containing adaptor protein). MAL is used by TLR4, TLR5, and TLR2 in combination

with TLR1 or TLR6. The alternative, 'TRIF (TIR domain-containing adaptor inducing interferon-β)-dependent pathway' is activated either by TRIF binding directly to TLR3, or indirectly to TLR4 with the help of the bridging adaptor TRAM (TRIF-related adaptor molecule). The relevance of MAL to human disease is highlighted by the effect of a number of SNPs in its gene on the susceptibility to infectious disease or autoimmunity (3).

The TIR domains typically comprise 125–200 residues with an alternating β-strand/α-helix arrangement of secondary structure elements, resulting in a central five-stranded parallel β-sheet (strands βA–βE) surrounded by five α-helices (αA–αE) on both sides of the sheet (2). The loops are usually named based on the elements of secondary structure they connect; for example, the 'BB loop' connects strand βB and helix αB. In addition to TLRs and TLR adaptors, TIR domains are found in interleukin-1 receptors, plant innate immune proteins and bacterial proteins (some of which interfere with the host innate immune responses) (2). TIR domains generally function through homotypic protein:protein interactions with other TIR domains. MAL contains a TIR domain at the C-terminus, and a short N-terminal region that contains a PIP2 (phosphatidylinositol 4,5-biphosphate)-binding motif and is responsible for its membrane localization (4).

The crystal structures of the TIR domain of MAL, as well as the crystal structures of two disease-associated mutants (S180L and D96N), have been reported independently by four groups (5-8). All of these structures consistently show unique features compared to other TIR domains. In particular, these structures lack helix αB and feature a long AB loop situated between the helix αA and the β-strand βB, which includes the conserved and functionally important BB-loop proline-containing motif (9-11). Furthermore, these crystal structures contain two disulfide bonds involving residues C89:C134 and C142:C174, respectively. The presence of these disulfides is unusual for a cytosolic protein and poses an intriguing possibility of redox control of MAL-mediated signalling (12).

To shed light on the role of cysteine residues in the structure and function of MAL, comprehensive biochemical and structural characterization of the protein was carried out. Here, data is presented showing that the recombinantly-produced protein does not contain disulfides, but several of its cysteines are highly reactive. The solution structure of the reduced form of the MAL TIR domain was determined (residues 79–221; MAL^{TIR}) by nuclear magnetic resonance (NMR) spectroscopy, revealing a more typical TIR-domain fold compared to the crystal structure, and uncovering a remarkable structural rearrangement upon disulfide formation that includes the relocation of a β-strand and repositioning of the BB-loop motif.

3.3 Methods

3.3.1 *Expression and purification of MAL^{TIR} and its cysteine mutants*

BL-21, SHuffle (New England Biolabs) or Origami (Merck Millipore) *Escherischia coli* cells were transformed by heat-shock with the pMCSG7 vector (13) encoding wild-type MAL^{TIR} (residues 79–221) or its mutants C89A, C91A, C116A, C134A, C142A, C157A and C174A. The cells were then grown in a 50 mL starter culture of LB media with 100 mg/L of ampicillin overnight at 37°C. Cells were centrifuged at 800 x *g* and washed in M9 salts. One mL of starter culture was added to each flask containing 500 mL auto-induction media (14), with ^{15}N-labeled ammonium chloride added for the expression of ^{15}N-labeled protein (15), and grown while shaking at 37°C until the OD of each flask reached an absorbance of 0.8 at the wavelength of 600 nm. Once the OD was reached, the temperature was reduced to 20°C for overnight protein expression. The cells were then lysed by sonication, and the protein was purified using nickel-affinity chromatography. TEV (tobacco etch virus) protease was added to the protein for overnight incubation at 4^{o}C to cleave the 6-histidine tag from the protein. The protein was further purified by re-eluting over the nickel affinity column, followed by size-exclusion chromatography with a Superdex 75 gel-filtration column (GE Healthcare 26/600 S75) into a buffer consisting of 20 mM Tris-HCl (pH 8.6) with 200 mM NaCl.

3.3.2 *Expression and purification of ^{15}N and ^{13}C-labelled MAL^{TIRC116A}*

The plasmid coding for the C116A mutant of MAL^{TIR} (MAL^{TIRC116A}) was transformed into *E. coli* BL-21 cells by heat-shock, and grown overnight in a starter culture of Luria's broth (LB) in the presence of 100 mg/L ampicillin while shaking at 37°C. The protein was then expressed in the LB media while shaking at 37°C until the cells reached an optical density (OD) of 0.7 at the wavelength of 600 nm. The sample was centrifuged at 800 x *g*, washed in M9 salts (Section 2.2.9) and resuspended in minimal M9 media (Section 2.2.9) containing ^{13}C glucose and ^{15}N ammonium chloride until the cell density reached an OD of 0.8 at the wavelength of 600 nm. The temperature was then reduced to 20°C and the cells were induced with 1 mM isopropyl-1-thiogalactopyrano-side (IPTG) for overnight expression. The protein was purified as described above.

3.3.3 *NMR data acquisition*

300 μL of $^{13}C/^{15}N$-labelled MAL^{TIR} sample containing 5% D_2O was added to a susceptibility-matched 5 mm outer-diameter microtube (Shigemi Inc., Japan). NMR data for wild-type MAL^{TIR} was acquired using a concentration of 168 μM of protein in Tris buffer at pH of 8.6 with 200 mM NaCl. A $MAL^{TIRC116A}$ mutant sample was prepared at a protein concentration of 600 μM of protein in TRIS buffer at pH of 8.6 with 200 mM NaCl.

NMR data were acquired at 18˚C using a 900 MHz AVANCE spectrometer (Bruker BioSpin, Germany) equipped with a cryogenically cooled probe. Data used for resonance assignment were acquired using non-uniform sampling (NUS); sampling schedules that approximated the rate of signal decay along the various indirect dimensions were generated using sched3D (16). The decay rates used were 1 Hz for all constant-time ^{15}N dimensions, 30 Hz for all ^{13}C dimensions, and 15 Hz for the semi-constant indirect ^{1}H dimension of the HBHA(CO)NH experiment. All pulse programs were executed in NUS mode on Topspin 3.2. ^{13}C- and ^{15}N-edited HSQC-NOESY experiments were acquired using linear sampling. Separate experiments were acquired for the aliphatic and aromatic regions of the ^{13}C dimension. The low concentration of wild-type MAL^{TIR} protein and the high pH and salt concentrations resulted in very poor signal-to-noise ratios and despite extensive signal averaging (128 scans per increment), the assignment rate using the HNCACB/CBCA(CO)NH, HNCO/HN(CA)CO and HNCA/HN(CO)CA sets of spectra did not yield high levels of assignments and further experiments for side-chain and NOE assignments were not undertaken.

$MAL^{TIRC116A}$ at the high concentration of 600 μM was found to remain mostly monomeric for ~18 hours, after which the signal intensities in the NMR spectra would broaden significantly due to self-association. The self-association could be reversed completely when the sample was incubated at 4˚C. The NMR data for this sample were therefore acquired in 16–20 hour segments. The HNCACB, HBHA(CO)NH, ^{15}N-HSQC-NOESY and ^{13}C-HSQC-NOESY (aliphatic region) were acquired as 3 sections, the CBCA(CO)NH, HN(CA)CO, H(C)CH and ^{13}C-HSQC-NOESY (aromatic region) as 2 sections and the HNCO as a single section (~18 days of acquisition time at 900 MHz). The segments were co-added in the time domain prior to reconstruction, either using traditional processing methods (linear prediction, apodisation, zerofilling followed by Fourier transform) or using automated maximum entropy reconstruction as described previously (17).

NMR spectra were analysed and assigned using the program CCPNMR (18). $^{1}H_N$, ^{15}N, $^{13}C_\alpha$, $^{13}C_\beta$, and $^{13}C'$ resonance assignments were obtained from analysis of amide-proton strips in 3D HNCA,

HN(CO)CA, HNCACB, CBCA(CO)NH, HN(CA)CO and HNCO spectra. Side-chain ^{1}H and ^{13}C chemical shifts were obtained primarily from a 3D H(C)CH-TOCSY and the remaining side-chain assignments were derived from 3D and ^{15}N- and ^{13}C-edited NOESY-HSQC spectra. Several segments (generally loop segments) of the protein could not be assigned due to either line broadening from conformational exchange or unfavourable NH exchange at the high pH required. The following sections/residues could not be assigned: 79–86, 100, 111, 125, 127, 131–133, 169, 170, 183, 184, 192–197, 201, 202, 206, 207. This resulted in an overall assignment rate of 72% of all ^{1}H atoms.

3.3.4 *Structure determination*

Distance restraints for structure calculations were derived from 3D ^{13}C- and ^{15}N-edited NOESY-HSQC spectra acquired with a mixing time of 90 ms. NOESY spectra were manually peak picked and integrated. The peak lists were then assigned and an ensemble of structures calculated automatically using the torsion angle dynamics package CYANA (19,20). The tolerances used for CANDID were 0.03 ppm in the indirect ^{1}H dimension, 0.02 ppm in the direct ^{1}H dimension and 0.4 ppm for the heteronuclei (^{13}C and ^{15}N).

Backbone dihedral-angle restraints (106 ϕ and 110 for ψ) were derived from TALOS+ chemical shift analysis (21,22); the restraint range was set to twice the estimated standard deviation. All X-Pro peptide bonds were clearly identified as *trans* on the basis of the C_β and C_γ chemical shifts for the Pro residues.

CYANA was used to calculate 200 structures from random starting conformations, then the 20 conformers with the lowest CYANA target function were chosen to represent the structural ensemble. During the automated NOESY assignment/structure calculation process, CYANA assigned ~80% of all NOESY crosspeaks (2648 out of 3398) for $MAL^{TIRC116A}$. The short mixing time minimises spin diffusion effects and the relatively large fraction of unassigned NOEs is consistent with the relatively high percentage of unassigned atoms (~28%).

3.3.5 *NMR of MAL^{TIR} at varying pH values*

To compare the effect of pH on the structure of MAL^{TIR}, the ^{15}N-labelled $MAL^{TIRC116A}$ was divided and buffer-exchanged into three buffers, each made of 10 mM HEPES and 50 mM NaCl, at pH 7.5, 8.0 or 8.6. The chemical shifts of ^{1}H and ^{15}N were measured at each pH value and analysed using the program CCPNMR (18).

3.3.6 *Determination of redox shifts of cysteines in wild-type MAL^{TIR}*

Stock solutions of reduced (GSH) and oxidised (GSSG) glutathione were made in 10 mM HEPES buffer and adjusted to pH 7.5. Nitrogen was bubbled through both solutions to remove oxygen. A ^{1}H spectrum of each stock was measured using a 900 MHz magnet and analysed using the spectrometer software (TopSpin 3.2) to confirm whether the stock solution redox species were indeed reduced or oxidised. Purified ^{15}N-labelled wild-type MAL^{TIR} in a 10 mM HEPES (pH 7.5) buffer was added to 5 different buffers containing varying ratios of GSH:GSSG that equate to specific electrochemical potentials (GSH only, -225 mV, -198.67 mV, -189.77 mV and GSSG only). Electrochemical potential was calculated using the Nernst equation $E° = E°_{GSH/GSSG} - (RT/nF) \ln(K_{eq})$, where $E°_{GSH/GSSG}$ is the standard potential of glutathione at pH 7.5 (-240 mV) (23), R is the universal gas constant (8.314 J K^{-1} m^{-1}), T is the absolute temperature, *n* is the number of electrons transferred, *F* is Faraday constant (9.648×10^{4} C mol^{-1}), and K_{eq} is the equilibrium constant ([GSH]^2/GSSG). A ^{1}H-^{15}N-HSQC was completed for each of the redox conditions at 298 K using a 900 MHz magnet to determine the chemical shifts of each residue.

3.3.7 *Analysis of chemical shifts of ^{1}H and ^{15}N atoms*

Changes in chemical shifts were quantified using the Euclidean method for distance measurements, following the frequency assignment of ^{1}H and ^{15}N atoms in each ^{1}H-^{15}N-HSQC experiment using the CCPNMR software (18). Individual residues were compared using d^2, where $d(p_1, p_2)^2 = (H_1 - H_2)^2 + (N_1 - N_2)^2$. Overall comparison between data-sets was calculated using root-mean-square deviation (RMSD); where RMSD = $\sqrt{\frac{1}{n}\sum d(p_1, p_2)}$.

3.3.8 *Reactivity of cysteine residues*

Intensities of peaks assigned to cysteine residues 89, 91, 116, 134, 142, 157 and 174 were measured in ^{1}H-^{15}N-HSQC spectra at redox potentials of -225, -199 and -190 mV using purified ^{15}N, ^{13}C wild-type MAL^{TIR}. To account for changes in peak intensities due to variations in sample conditions, such as the salt concentration, which impact on magnetic susceptibility, conductivity, and correlation time, the intensities of 12 peaks corresponding to atoms at coordinates distal to the cysteine residues in the 3D structure were monitored. Intensity changes of these unreactive sites were averaged and applied to the measured cysteine residue intensities as a normalisation factor.

3.4 Results

3.4.1 *Secondary structure of MAL^{TIR} in solution is different to that in the crystal structure*

The amide-backbone of purified ^{13}C and ^{15}N-labelled wild-type MAL^{TIR} provided limited data due to its inherent insolubility as discussed in Chapter 2. Secondary structure prediction of the limited chemical shifts obtained from the wild-type MAL^{TIR} using TALOS+ indicated that there is a major structural rearrangement between residues 111 and 138, compared to that of the crystal structure (Figure 3.1). The crystal structures indicated that this region of the protein formed a long loop, called AB-loop, with high flexibility (5-8). The predicted solution structure suggested the formation of a Bβ-strand followed by a BB loop, which is a key feature in typical TIR domain folds (2). This result prompted further investigation to find a mutant of MAL^{TIR} that would be soluble for the length of the experiment, to gain better quality data for structural analysis by NMR.

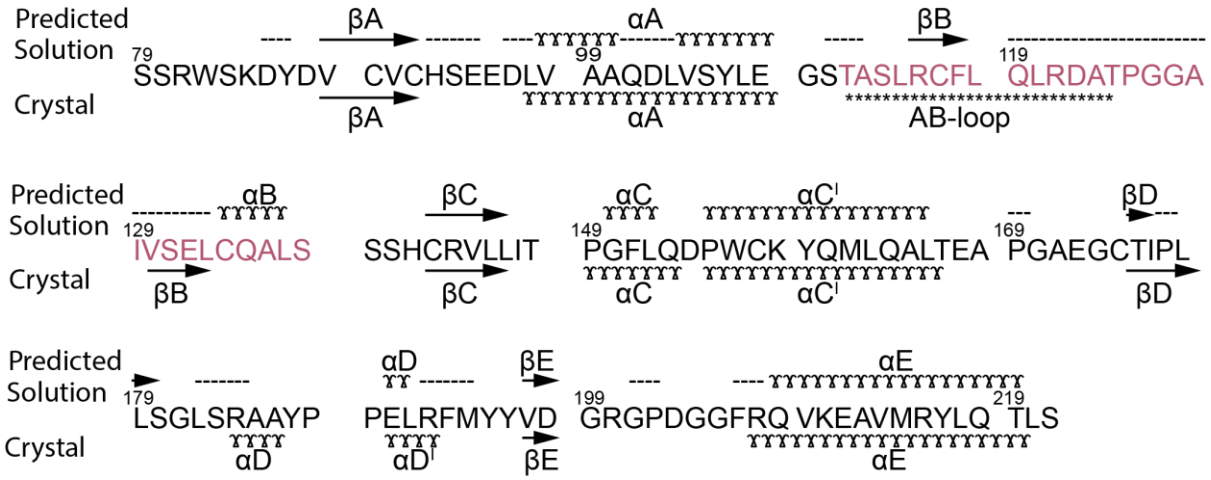


Figure 3.1 Comparison of the predicted secondary structure of wild-type MAL^{TIR} in solution from amide back-bone chemical shifts using TALOS+, and crystal structure. (--) missing chemical shifts, (ϒϒ) α-helix, (⟶) β-strand. Red residues indicate the region with a major structural rearrangement between the two structures.

3.4.2 *Single cysteine mutations have varying impacts on the MAL structure*

Purified wild-type MAL^{TIR} provided limited data due to its tendency to precipitate in solution at high concentrations over a span of 12 hours at 18°C. To study the structural features of MAL^{TIR}, cysteine-to-alanine mutants were explored to find a protein that retained the wild-type fold while remaining soluble for the length of the experiment. Overall, most cysteine-to-alanine mutants of MAL^{TIR} retained a structure similar to the wild-type protein, as demonstrated by minor chemical shift

differences between the atoms of mutants and the wild-type protein (Table 3.1). The RMSDs for the cysteine mutants compared to the wild-type protein ranged from 0.028 to 0.120 ppm, with C89A having the lowest perturbation and C157A the highest. Neighbouring residues were not included in the calculation, to minimise the local effect of the mutations on the chemical shifts of the protein. The mutant C157A displayed a consistently high ^{1}H and ^{15}N displacement, with eight residues throughout the protein backbone containing shifts over 0.2 ppm (Appendix Figure 3.6). By contrast, the C116A mutant had an RMSD of 0.058 ppm compared to the wild-type protein and contained only three ^{1}H-^{15}N shifts over 0.2 ppm. Two of these are either neighbouring (R115; shift of 0.46 ppm) or in close spatial proximity (D87; shift of 0.227 ppm) to the mutant residue. The mutants C89A, C134A and C142A contained no ^{1}H-^{15}N shifts over 0.2 ppm, with RMSD values of 0.033, 0.028 and 0.042 ppm, respectively, when compared to the wild-type protein, while C91A and C174A each contained one residue over 0.2 ppm, but maintained low RMSD values of 0.050 and 0.045 ppm, respectively, when compared to the wild-type protein.

Solubility of MAL^{TIR} mutants was analysed by measuring a ^{1}H-^{15}N-HSQC spectrum at day zero and again at day six following incubation at 25°C. Mutants C89A and C116A remained soluble over this time period while the rest of the mutants formed precipitates in solution. ^{1}H-^{15}N-HSQC analysis revealed that C116A produced very minor differences following a 6 day incubation at 25°C. Given that C116A retained similar fold to the wild-type (Figure 3.2), remained soluble for a 6 day period (Figure 3.3), expressed well (section 2.2.3), and did not involve any disulphide bonds in the crystal structure that would bias this structure away from the crystal structure, this mutant was chosen for structural studies.

Table 3.1 RMSD values of cysteine-to-alanine mutants of MAL^{TIR} compared to the wild-type protein, calculated using Eucladian norm.

Cysteine mutant	RMSD (ppm)
C89A	0.033
C91A	0.050
C116A	0.058
C134A	0.028
C142A	0.042
C157A	0.120
C174A	0.045

See **Appendix Figures 3.1–3.7** for residue-by-residue chemical shift comparison to the wild-type protein. Neighbouring residues of the mutation were not included in this calculation to minimise local effects on the protein.

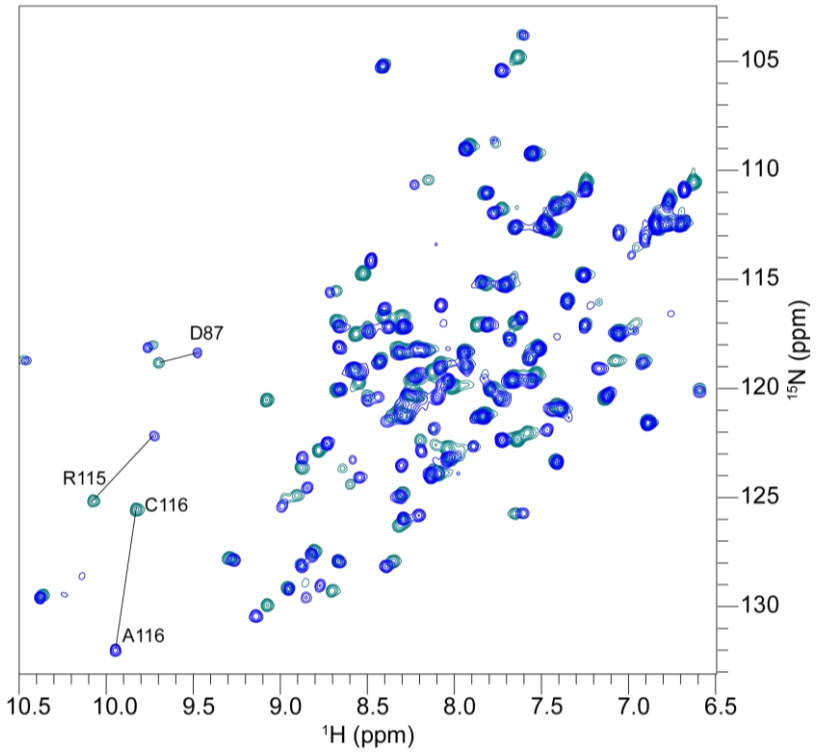


Figure 3.2 Overlay of ^{1}H-^{15}N-HSQC spectra from MAL^{TIRC116A} (blue) and wild-type MAL^{TIR} (cyan). Most chemical shifts of ^{1}H and ^{15}N are conserved between the spectra, with a large shift detected for the mutated C116 residue, the neighbouring residue R115 and additionally, residue D87.

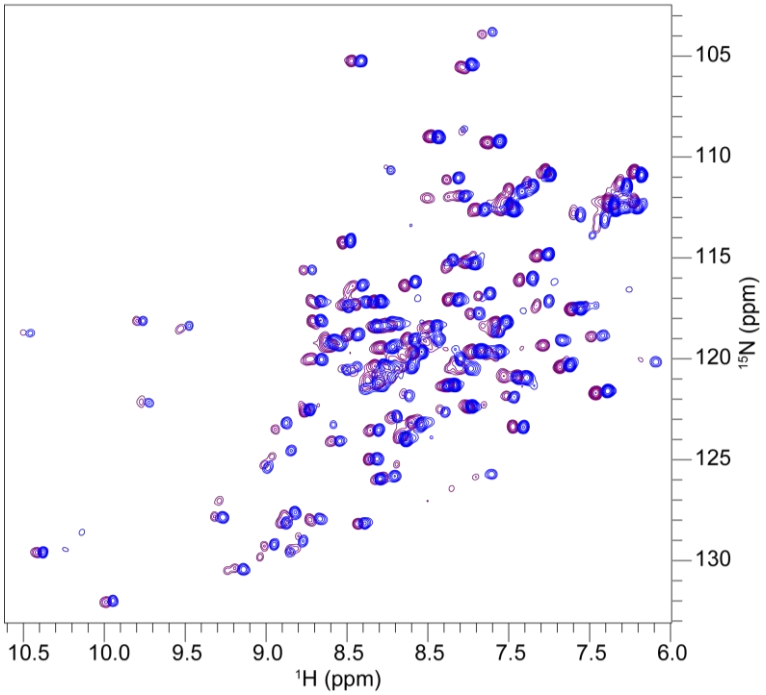


Figure 3.3 Overlay of ^{1}H-^{15}N-HSQC spectra of MAL^{TIRC116A} measured at 18°C (blue) at time 0, and after 6 days, measured at 25°C (purple). The overlay of spectra shows that the fold of the protein has been retained. The spectra have been shifted by 0.1 ppm in the ^{1}H axis to aid comparison.

3.4.3 *Chemical shifts of the cysteine residues*

The chemical shift of the carbon-β atom of the cysteine residues in MAL^{TIRC116A} mutant was measured using the CBCA(CO)NH experiment, to define their redox state following expression in BL-21 *E.coli* cells. The chemical shift of the carbon-β atom that is less than 32.5 ppm has previously been defined to constitute a reduced cysteine residue (24,25). The chemical shift of six cysteine residues in MAL^{TIRC116A} were consistent with the reduced state with a chemical shift of less than 32.5 ppm (Table 3.2). This result is consistent with that observed in the wild-type protein (Section 2.2.12) suggesting that the redox state of MAL^{C116ATIR} mutant is the same as that of the wild-type protein.

Table 3.2 Cysteine carbon-β shifts of $MAL^{TIRC116A}$

Cysteine residue	Carbon-β shift (ppm)	Redox state (<32.5 ppm = reduced)
89	28.8	Reduced
91	28.3	Reduced
116	-	-
134	29.5	Reduced
142	31.4	Reduced
157	25.8	Reduced
174	28.6	Reduced

3.4.4 *NMR solution structure of $MAL^{TIRC116A}$ reveals remarkable structural differences from the crystal structures*

The lowest energy solution structure of human $MAL^{TIRC116A}$ is shown in Figure 3.4 A (PDB 2NDH). The top-twenty lowest overall energy structures generated by calculations using TALOS+ for backbone dihedral angle restrains (21,22) and CYANA software for torsion angle dynamics calculations (19,20) yielded a heavy-atom amide-backbone (NH, C_α, C') RMSD of 1.28 Å (Table 3.3).

Typical TIR-domain structures feature a central five-stranded parallel β-sheet (strands βA–βE) surrounded by five α-helices (αA–αE) on both sides of the sheet (2). By contrast, the crystal structures of MAL^{TIR} (5-8) revealed the absence of helix αB; the sequence corresponding to strand βB in typical TIR domain structures is instead located in the long and flexible AB loop. The solution structure of the reduced form of $MAL^{TIRC116A}$ reveals that the secondary-structure arrangement is more consistent with a typical TIR-domain fold than with the crystal structures of MAL^{TIR}. The AB loop connecting αA and βB is short, comprising only 7 residues (109–115). This loop is followed by the strand βB (residues 116–119), as supported by a large number of short to medium-range inter-residue NOEs (Figure 3.6). This arrangement allows for the formation of a long flexible BB-loop between the βB strand and the helix αB; the loop corresponds to residues 120–134, which contain many of the unassigned atoms, and where assigned, show few long-range NOEs, suggesting inherent flexibility.

The $MAL^{TIRC116A}$ mutant contains six cysteine residues. Consistently with the wild-type protein, the chemical shifts of these six cysteine C_β atoms range from 25.8 to 31.4 ppm, which indicates that all cysteine residues are reduced. The solution structure contains one pair of cysteine residues (C142

and C174) in close-enough proximity (5.4–6.9 Å) for possible disulfide-bond formation; however, chemical shifts of cysteine C_β atoms indicate that they are both reduced.

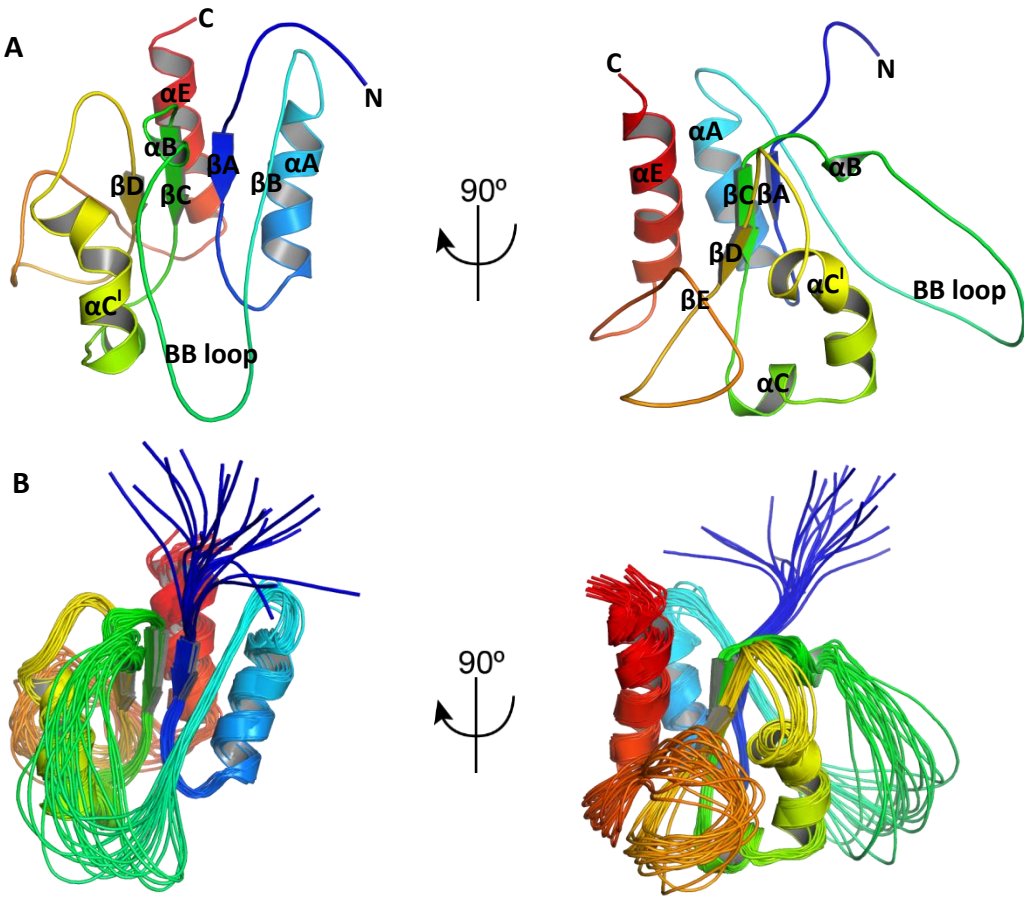


Figure 3.4 A) NMR solution structure of human MAL^{TIRC116A} (PDB ID: 2NDH). **B**) Top-twenty lowest-energy models, showing the inherent flexibility of the BB-loop in green.

Table 3.3 NMR structure statistics of MAL$^{TIRC116A\ (a)}$

Experimental restraints[b]		
	Inter-proton distance restraints	
	Intra-residue	362
	Sequential	422
	Medium-range (i–j < 5)	317
	Long-range (i–j > 5)	262
	Dihedral-angle restraints	216
	Total number of restraints per residue	11.12
RMSD from mean coordinate structure (Å)[c]		
	Backbone atoms (residues 85–120, 134–179 & 204–220)	0.94 ± 0.17
	All heavy atoms (residues 85–120, 134–179 & 204–220)	1.28 ± 0.14
Stereochemical quality[d]		
	Residues in most favoured Ramachandran region (%)	78.0
	Ramachandran outliers (%)	0 ± 0

[a]All statistics are given as mean ± SD.

[b]Only structurally relevant restraints, as defined by CYANA, are included.

[c]Mean r.m.s. deviation calculated over the entire ensemble of 20 structures.

[d]As reported by CYANA (19,20)

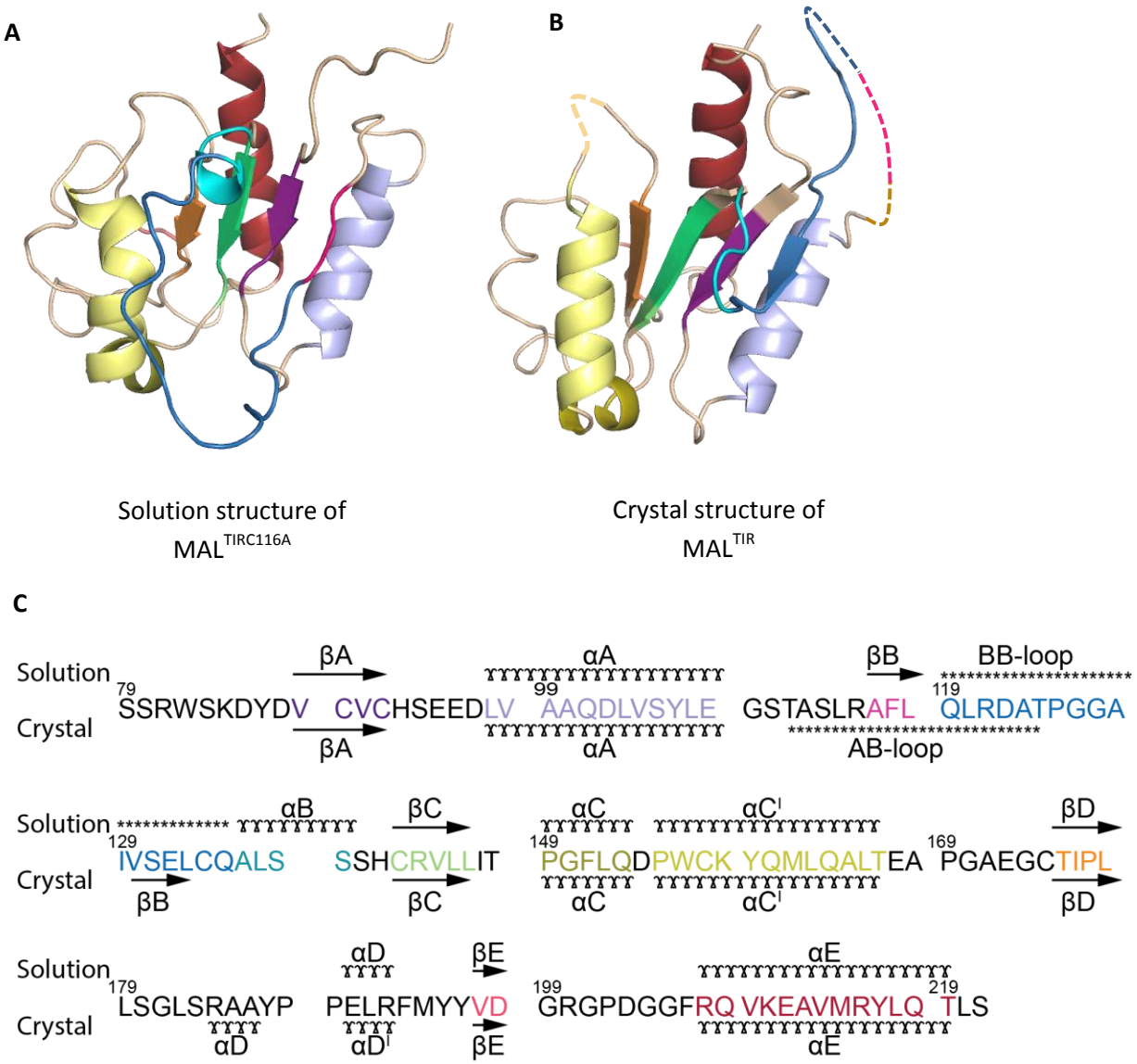


Figure 3.5 Comparison of the solution and crystal structures of MAL^{TIR}. Specific regions are coloured to highlight the structural re-arrangements between the two structures. (A) Solution structure of MAL^{TIR}, shown in cartoon representation. (B) Crystal structure of (5) (PDB ID 2Y92). Regions in the crystal structure comprising residues 112–123 (AB loop) and 168–171 did not have interpretable electron density and were not included in the crystal structure model; they are shown here for illustration purposes based on *ab-initio* modelling and energy-minimisation (5). (C) Structure-based sequence alignment, showing the elements of secondary structure. The solution secondary structure is based on the calculated structure using CYANA (19,20). (ϒϒ) α-helix, (⟶) β-strand, (****) loop.

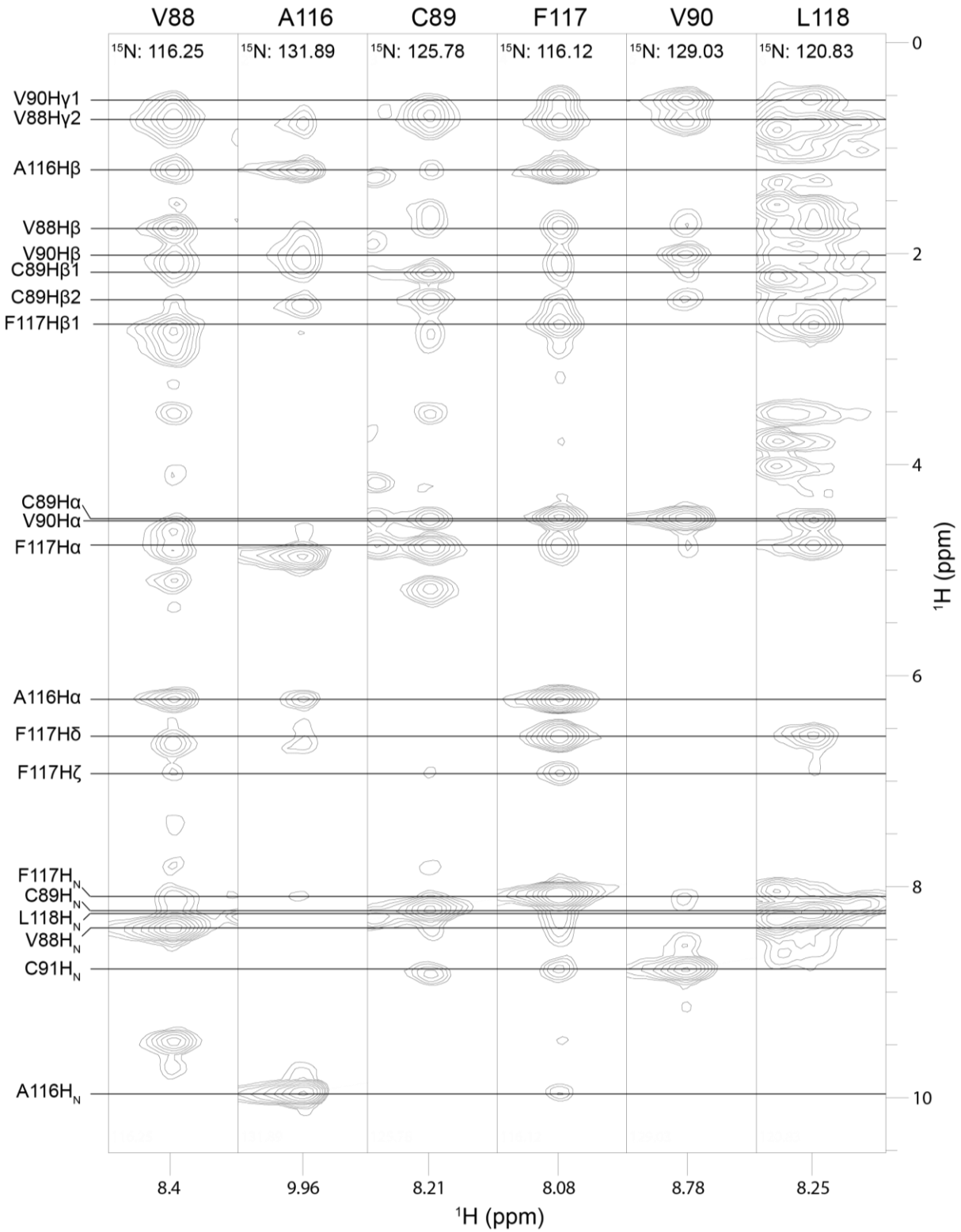


Figure 3.6 ^{15}N-NOESY strips from MAL^{TIRC116A} exhibiting short to medium-ranged NOE cross-peaks between residues situated in the βA strand (V88, C89, V90) and the βB strand (A116, F117,L118). The residues that participate in hydrogen-bond pairs to form the β-sheet are: V88-A116, C89-F117, and V90-L118. The large number of NOEs detected between theses pairs of residues indicates that these residues are in close proximity to one another.

3.4.5 *MAL^{TIR} retains its structure at physiological pH*

The chemical shifts of the ^{15}N and ^{1}H atoms from MAL^{TIR} were compared after the protein was buffered at a pH of 7.5, 8.0 or 8.6. Overall, the chemical shifts remained similar (Figure 3.7), suggesting that the structure remains unchanged over the tested pH conditions. An RMSD of 0.017 ppm was calculated between the spectra at pH 7.5 and 8.6. Residue L120 underwent the largest shift when exposed to pH 8.6, with shifts of 0.081 ppm when compared to pH 7.5, while other residues underwent minor changes (Appendix Figure 3.8). This change in chemical shift may be due to the proximity of the NH of L120 to the aromatic side-chain of the H92 residue, which is susceptible to protonation. Exposing the protein to pH 8.0 and 7.5 resulted in a total 9 and 12 new NH resonances, respectively, when compared to pH 8.6, consistent with the reduced NH exchange rate at the lower pHs.

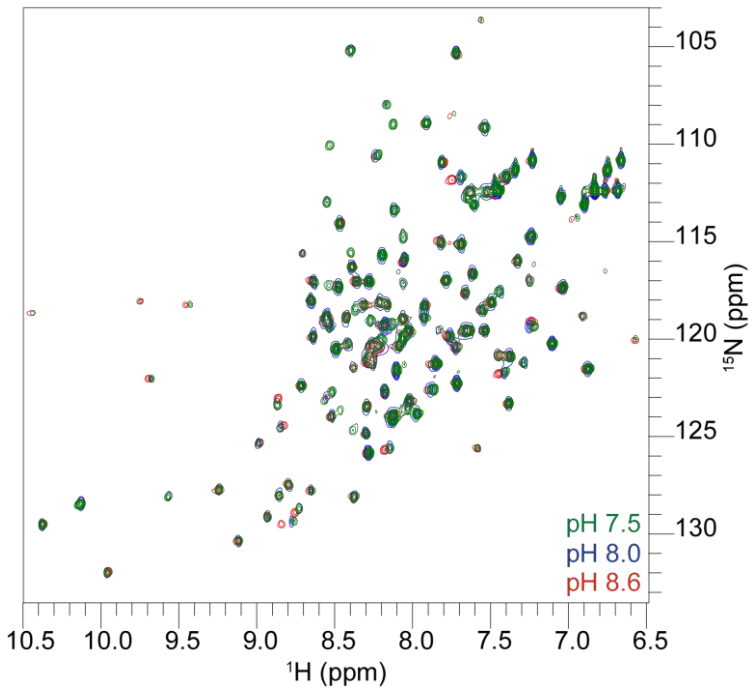
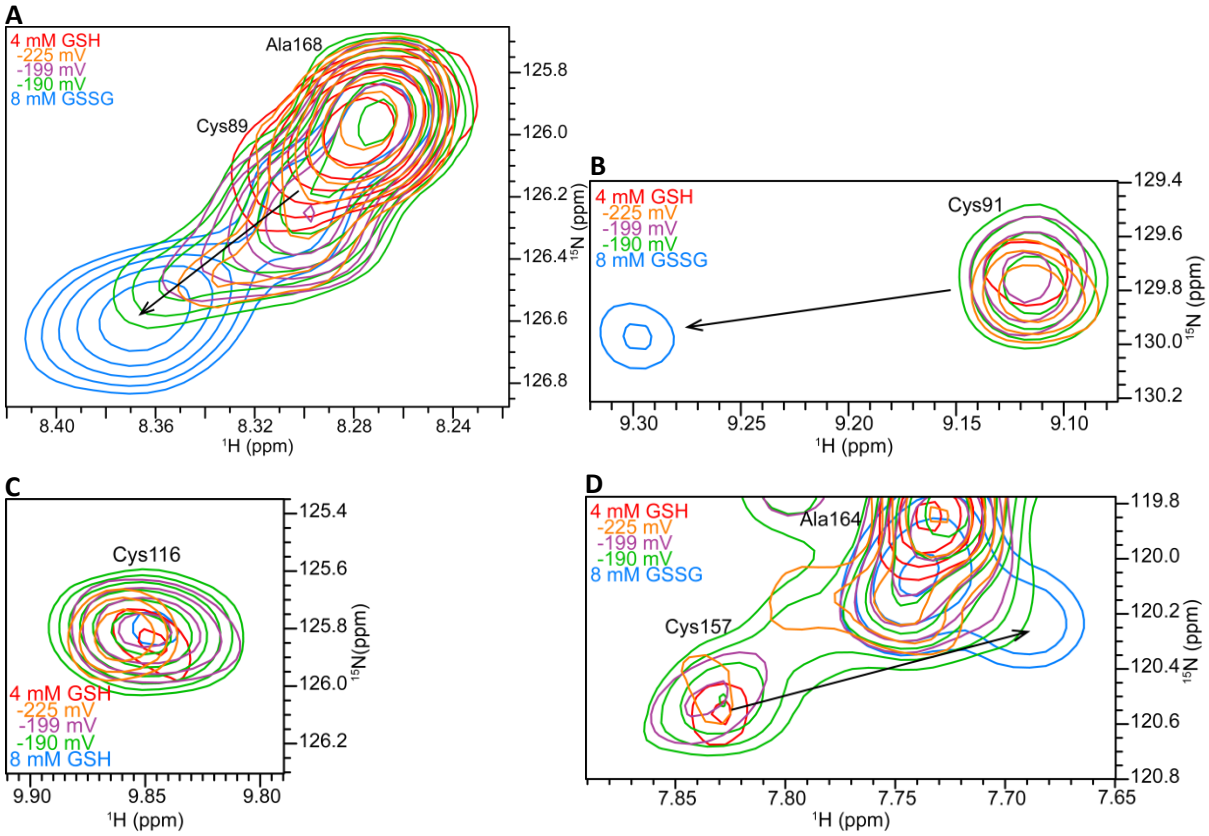


Figure 3.7 Overlay of ^{1}H-^{15}N-HSQC spectra of MAL^{TIR} at pH 7.5, 8.0 and 8.6.

3.4.6 *Cysteine reactivity at changing redox conditions*

Exposure of wild-type MAL^{TIR} to redox conditions that ranged from fully reducing to fully oxidising, with three physiological conditions (26) including -225, -199 and -190 mV, indicated that cysteines varied in reactivity. Major changes in chemical shifts were observed at the fully oxidising condition containing only GSSG, especially for cysteine residues 89, 91 and 157 (Figure 3.8). Using the ^{1}H-^{15}N-HSQC experiments that corresponded to the three redox-buffered physiological conditions (-225, -

199, -190 mV), changes of peak intensity belonging to each cysteine were measured and standardised against the highest intensity spectrum (-225 mV) (Figure 3.9). Cysteine residues C116 and 157 underwent almost no change with around 7% signal change at -190 when compared to -225 mV, indicating little oxidation of these cysteines at physiological redox potentials. Cysteines 89, 134, 142 and 174 underwent minor oxidation as indicated by small changes in signal intensities ranging from 15-29% at -190 mV compared to -225 mV. By contrast, cysteine residue 91 was the most reactive with 45% signal loss at -190 mV, compared to -225 mV, indicating that the major population of this residue had been oxidised.



Figures 3.8 Cysteine residues C89 (**A**), C91 (**B**), C116 (**C**), and 157 (**D**), showing ^{1}H and ^{15}N shifts following exposure to increasingly oxidising conditions.

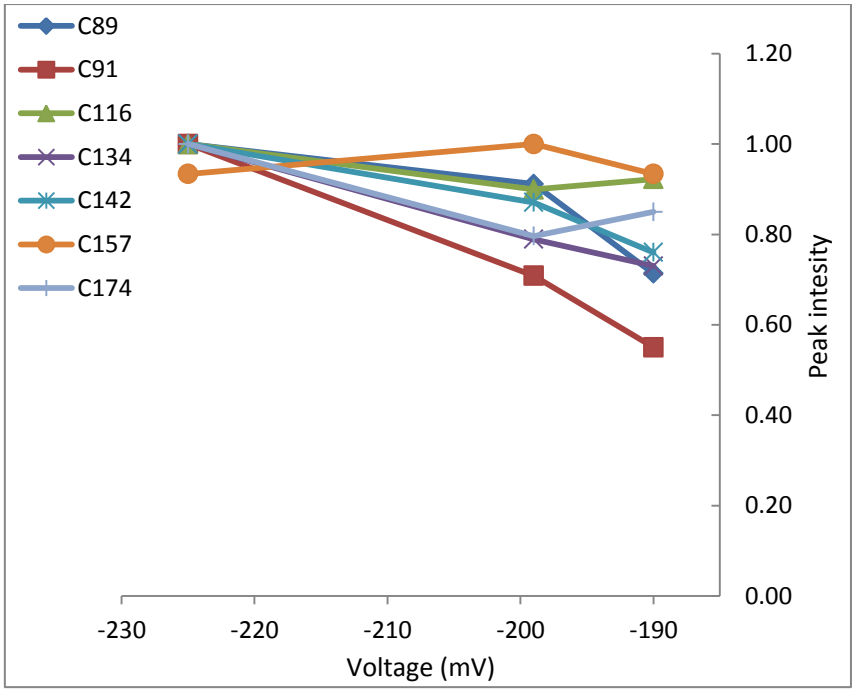


Figure 3.9 Reactivity of MAL cysteine residues at increasing redox potential.

3.5 Discussion

Based on the available MAL crystal structures, the construct comprising residues 79–221 (designated here MAL^{TIR}) was selected for structural studies by solution NMR spectroscopy. The C116A mutant ($MAL^{TIRC116A}$) was found to have superior solubility and behaved as a stable monomer, compared with the wild-type and other cysteine-to-alanine mutants. The $MAL^{TIRC116A}$ structure was found to be representative of the wild-type MAL^{TIR} based on the comparison of chemical shifts and secondary structure prediction. The structure calculations of $MAL^{TIRC116A}$ produced an RMSD of 1.28 Å for the ensemble of the top 20 lowest energy states using the heavy atoms of the amide backbone (NH, C_α and C'). The structure was solved at a pH of 8.6 in a high ionic strength of 200 mM sodium chloride. However, the comparison of spectra of ^{15}N-labelled MAL^{TIR} in the pH range 8.6–7.5 shows that the structure is representative of a more physiological pH. The extreme conditions of high salt and pH yielded suitable sample concentrations for NMR analysis. These conditions, however, result in loss of signals at rapidly exchanging backbone amide NHs and poor signal-to-noise ratio of the NMR resonances. Despite these difficulties, an unconventional data acquisition strategy was used, and advanced data acquisition and processing methods at ultra-high fields were used to produce the first solution structure of MAL, be it at medium-low resolution.

Analysis of the amide-backbone chemical shifts (NH, C_α, C_β and C') showed a structural re-arrangement in the protein compared to the crystal structures (Figure 3.5). Consistent with other

'typical' TIR-domain structures, the solution structure contains a long BB-loop (residues 120–134) situated between the strand βB and the helix αB. The BB-loop is a conserved feature among TIR domains and a number of studies have identified it to be important for mediating homotypic TIR-domain interactions (9,11,27,28). The striking structural differences correspond to the secondary structure arrangement between residues 111 and 139 (Figure 3.5 C). The crystal structures contain a flexible AB-loop corresponding to residues 111–124, linking the helix αA and strand βB. This rearrangement is caused by cysteine residues 89 and 134 forming a disulfide bond and shifting the location of strand βB. In the solution structure, the C_α atoms of these two cysteine residues are ~11 Å apart, making disulfide-bond formation unlikely (29). The second disulfide bond found in the crystal structures involves the pair C142 and C174. In the solution structure, these residues remain in close proximity (~6 Å); however, the chemical shifts of the C_β atoms from these residues suggest that the cysteine side-chains are in a reduced state. In addition, these two cysteine residues are not conserved and are not likely required for the folding of the protein (5). Indeed, chemical shifts of the C_β atoms of all cysteine residues in the solution structure are consistent with a reduced state (24).

$MAL^{TIRC116A}$ was purified using an *E. coli* bacterial expression system. Similarly to mammalian cells, *E. coli* BL-21 cells also contain a reducing cell cytosol (30). The fact that the $MAL^{TIRC116A}$ solution structure presented here was measured to contain chemical shifts that are consistent with reduced cysteine residues is not surprising. Supporting evidence from mass spectrometry using wild-type MAL^{TIR} determined the presence of a fully reduced species (Section 2.2.6). Given that the mammalian cell cytosol is highly reducing, it is most likely that this structure represents the physiological state. The redox potential of the mammalian cytosol is fluid and undergoes change from about -225 mV to -190 mV, where -225 mV represents the cell proliferation stage and -190 mV when the cell is undergoing apoptosis (26). Under fully reducing conditions, there are minor changes in the chemical shift of the protein, whereas exposure to the fully oxidised condition results in substantial changes (Appendix Figure 3.9). The changes detected are most likely due to variations in the redox potential of the buffer condition. As the buffer changes from reducing to oxidising, large chemical shifts are detected in the cysteine residues that are likely to be the result of the oxidation state of the cysteine. The chemical shifts of cysteine residues in $MAL^{TIRC116A}$ measured in the ^{1}H-^{15}N-HSQC experiment show that they are all reduced at -225 mV. Assessment of cysteine reactivity revealed that residue 91 was the most sensitive to oxidation, whereby peak intensity decreased by 45%, which is attributed to oxidation at a redox potential of -190 mV. In contrast, cysteine 116 was the least sensitive with no major changes detected, irrespective of the redox environment. The reactivity of C91 demonstrated here could explain the tendency to react with a DTT molecule as

found in crystal structures. Similarly, peak intensities of cysteine residues C89 and C134 decreased by ~28%, suggesting that they underwent oxidation in a -190 mV glutathione buffer. These two residues form a disulfide bond in all four crystal structures (5-8). This could explain the formation of the disulfide bond observed in the MAL^{TIR} crystal structures, and in turn, the structural rearrangement of the strand βB. This knock-on effect from this disulfide bond would result in the shortening of the BB-loop and the extension of the AB loop.

Out of the seven cysteine residues in MAL^{TIR}, only one (C157) is highly conserved (5). Amongst other TIR domains, the position equivalent to cysteine 89 in MAL corresponds to a conserved phenylalanine. Considering that the TIR domain fold is conserved, it is highly likely that the disulfide bond between residues C89 and C134 in the MAL^{TIR} crystal has formed over time required for crystallisation and the lack of sufficiently reducing environment.

Two of the four crystal structures showed the residues C157 and C91 coordinating a DTT molecule (5,6). In another crystal structure, these two cysteines were replaced by seleno-methionine (7), while a fourth structure showed no interactions with any buffer components (8). Although C157 is not as reactive as C91, data presented here shows that the right condition, such as exclusive exposure to GSSG, results in a significant chemical shift, indicating oxidation of this residue. In addition, mutating the highly conserved C157 residue to an alanine results in a large perturbation of the structure, suggesting that this residue is structurally important. Studies have shown that mutating residues in the helix αC, on which the conserved C157 resides, results in the inability of MAL to immunoprecipitate its binding partner MyD88 (5) and attenuate its ability to signal (6). This, together with the ability of C91 and 157 to be modified by glutathionlylation (Section 2.2.7), could indicate that these residues are an important pair that could have functional roles in the protein.

Increasing evidence suggests that during TLR4-mediated immune activation, the cytosolic environment becomes more oxidising (31). This can lead to the modification of cysteine residues on cytosolic proteins, most commonly to their oxidation by hydrogen peroxide species produced by NADPH (32), or by glutathionylation (33). The glutathione *S*-transferase omega class 1 (GSTO1-1) protein has been associated with the glutathionylation of cysteines, resulting in an on/off switch (34,35). The precise target protein of GSTO1-1 is unclear; however, there is strong evidence to suggest that it regulates proteins early in the TLR4 signalling cascade, which could involve the cytosolic TIR domains of TLR4, MAL or MyD88 (36). Taken together with the redox data presented here, the identification of redox sensitive cysteines, especially C91, would suggest that glutathionylation by GSTO1-1 of these MAL^{TIR} cysteines may play an important functional role.

Single nucleotide polymorphisms in MAL have a significant impact on the immune response of infected individuals. For example, a rare D96N mutation prevents MAL from binding to the downstream binding partner MyD88, resulting in no activation of NF-κB (37). Lin *et al* (6) similarly demonstrated that mutating residues that reside between R115 and E132 results in decreased NF-κB activation (6). Put into the context of the solution structure presented here, these mutations are located on the βB strand and BB-loop. This work has striking implications because these regions differ remarkably between the solution and crystal structures. The NMR solution model therefore provides a new framework for future work on the effects of mutations and modelling studies.

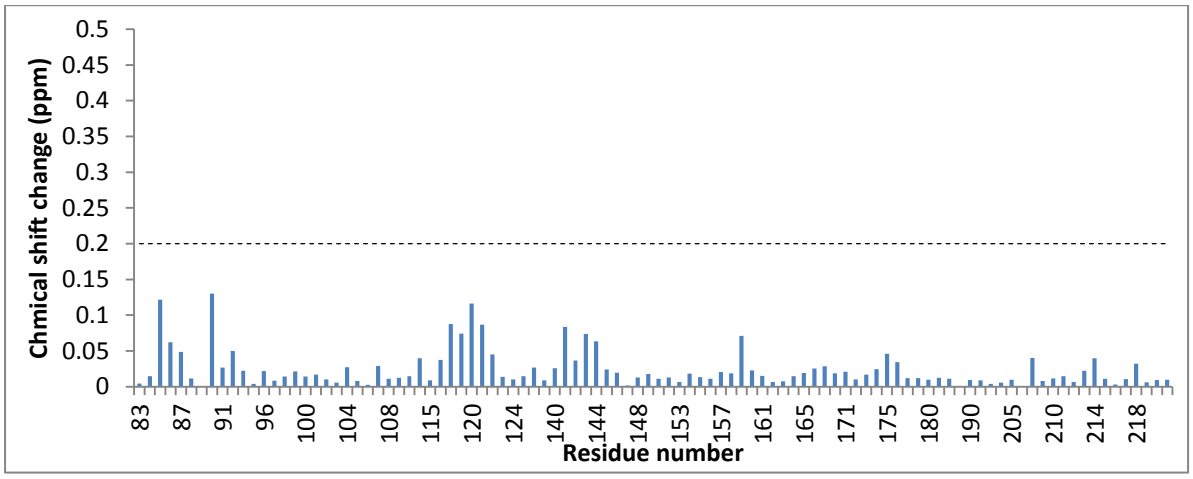
3.6 Conclusions

Results presented here show the solution structure of $MAL^{TIRC116A}$ in the reduced form, as determined by NMR spectroscopy. The structure shows striking differences to the crystal structures of MAL^{TIR} determined previously (5-8), which represent the oxidised form of the protein. Three of the seven cysteine residues in the reduced form are sensitive to redox change when exposed to oxidising conditions, for example through the addition of GSSG. The work suggests that the cysteines 91 and 157 could play an important role in the signalling and regulation involving the mammalian MAL adaptor protein. The crystal structure of MAL^{TIR} has been used in a number of attempts to model the interactions involving this adaptor during signalling (12,38-40). The structure presented here suggests that the crystal structure may not be the signalling-competent form. This new structure now provides a better template for modelling approaches.

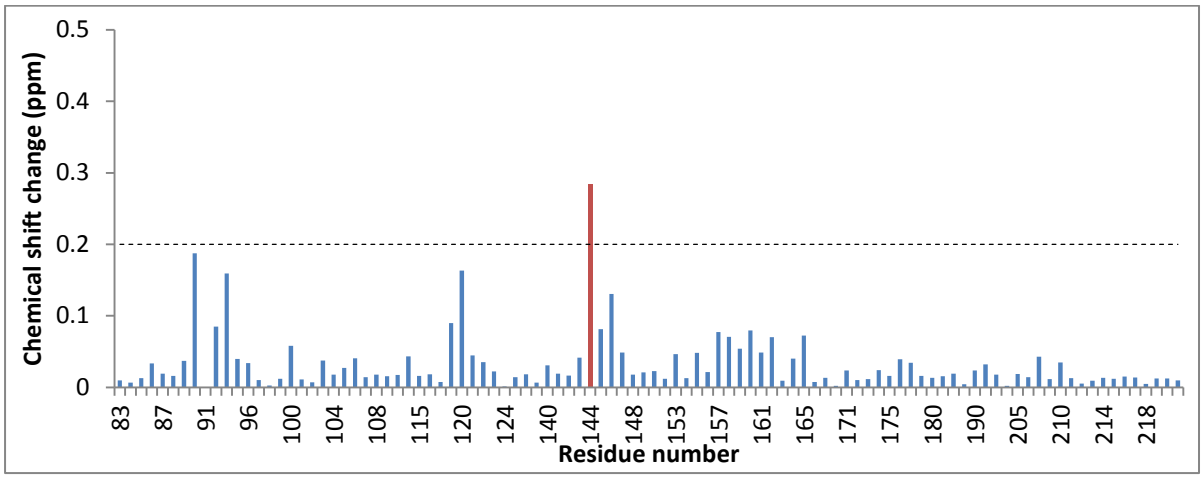
3.7 Chapter 3 appendix

Comparison of chemical shifts between the cysteine mutants and the wild-type protein (Appendix Figures 3.1–3.7)

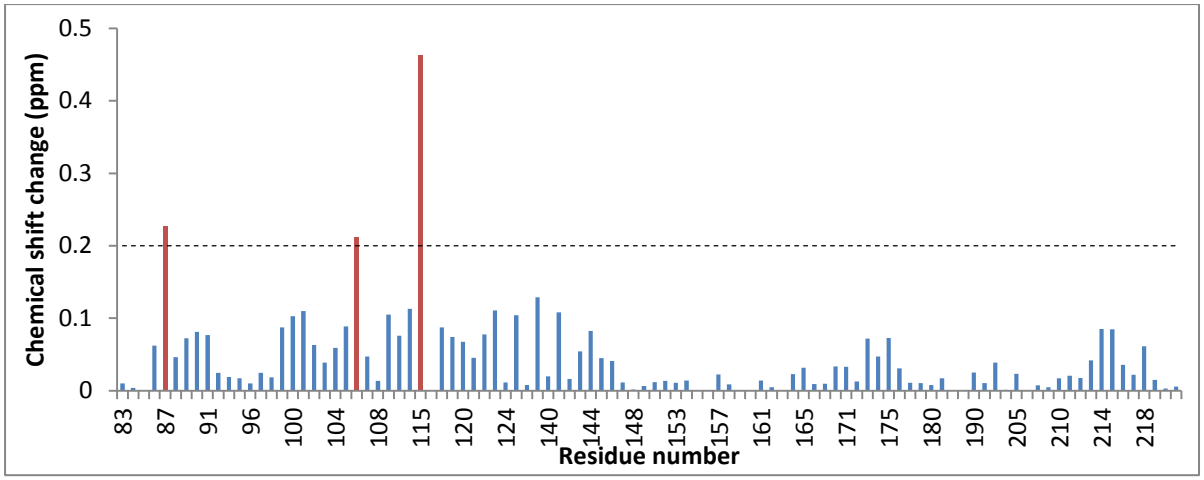
^{1}H and ^{15}N chemical shifts of residues from the MAL^{TIR} mutants are compared to the wild-type protein. An arbitrary threshold of 0.2 ppm is selected as a significant chemical shift change between the mutants and the wild-type protein and is indicated by a dashed line.



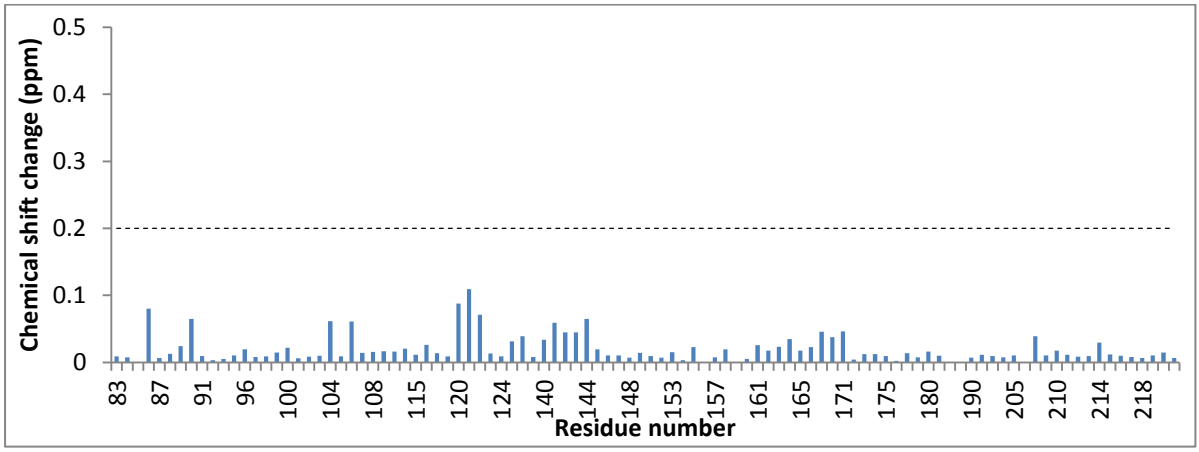
Appendix Figure 3.1 Comparison of chemical shifts between the C89A mutant and the wild-type MAL^{TIR} protein.



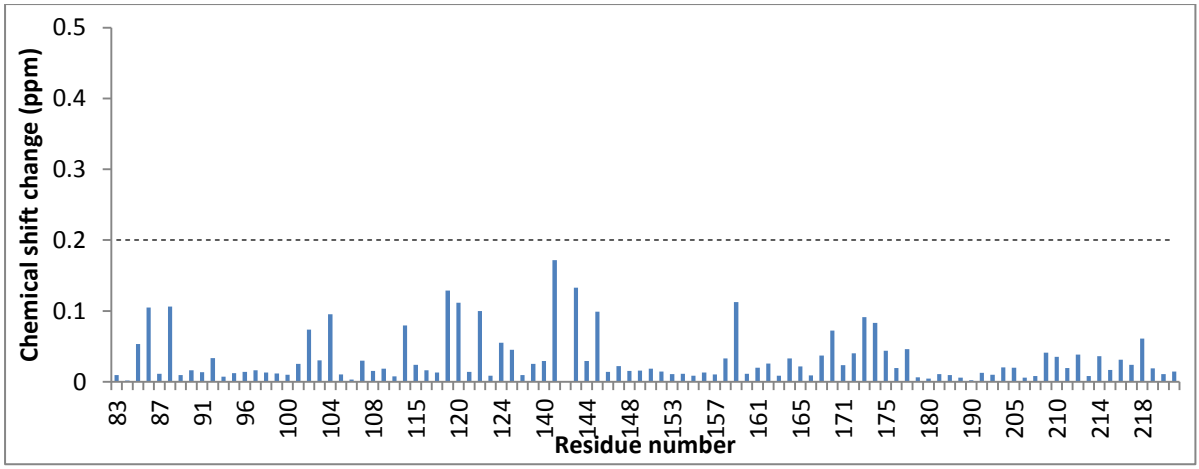
Appendix Figure 3.2 Comparison of chemical shifts between C91A mutant and the wild-type MAL^{TIR} protein.



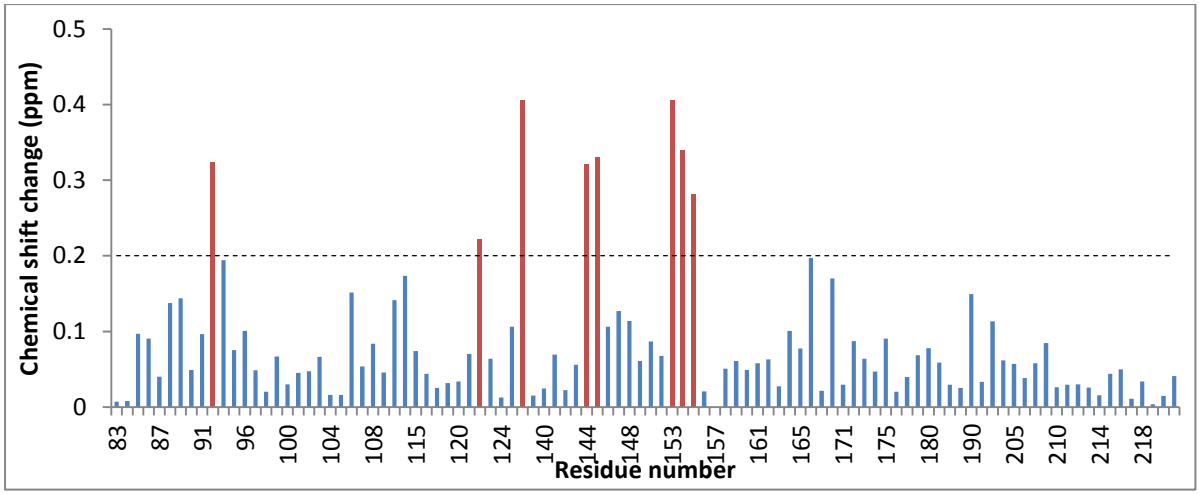
Appendix Figure 3.3 Comparison of chemical shifts between C116A mutant and the wild-type MAL^{TIR} protein.



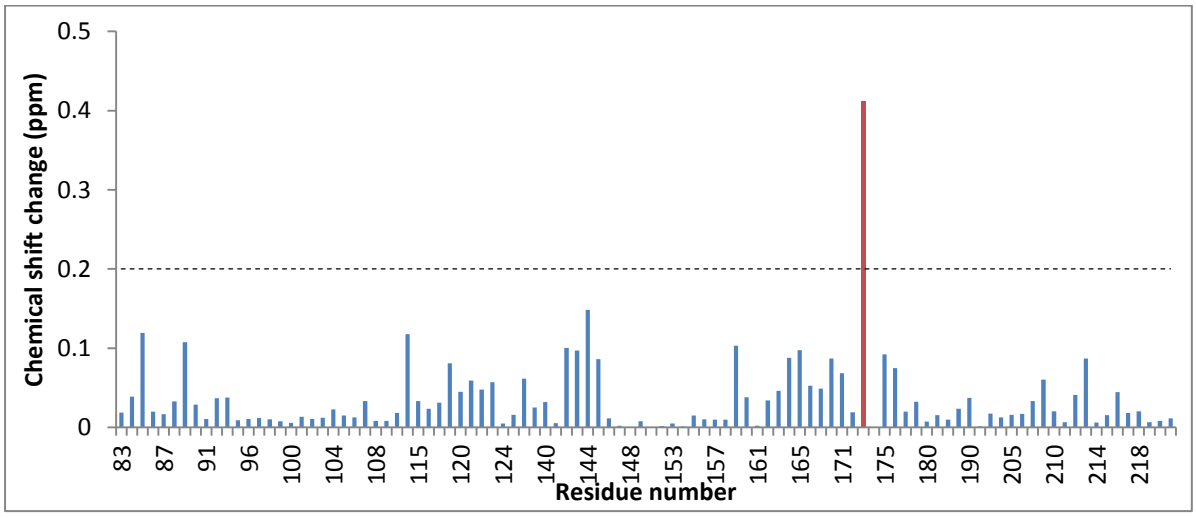
Appendix Figure 3.4 Comparison of chemical shifts between C134A mutant and the wild-type MAL^{TIR} protein.



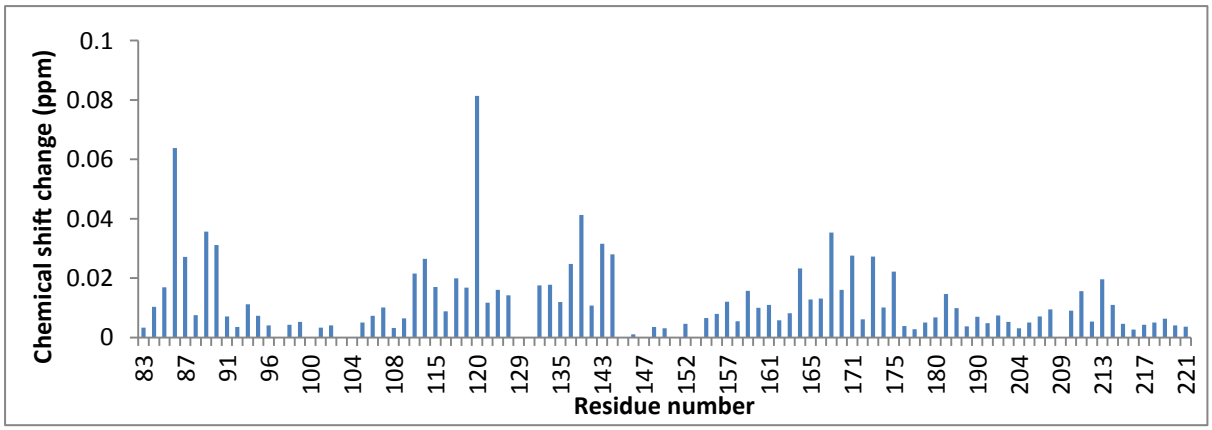
Appendix Figure 3.5 Comparison of chemical shifts between C142A mutant and the wild-type MAL^{TIR} protein.



Appendix Figure 3.6 Comparison of chemical shifts between C157A mutant and the wild-type MAL^{TIR} protein.



Appendix Figure 3.7 Comparison of chemical shifts between C174A mutant and the wild-type MAL^{TIR} protein.



Appendix Figure 3.8 Comparison of chemical shifts for MAL^{TIR} at pH 7.5 and 8.6.

Appendix Figure 3.9 RMSD comparison of ^{1}H and ^{15}N chemical shifts between different redox conditions.

Condition	GSH only
-225 mV	0.019
-199 mV	0.019
-190 mV	0.018
GSSG only	0.049

3.8 Chapter 3 references

1. Gay, N. J., Symmons, M. F., Gangloff, M., and Bryant, C. E. (2014) Assembly and localization of Toll-like receptor signalling complexes. *Nature Reviews Immunology* **14**, 546–558.
2. Ve, T., Williams, S. J., and Kobe, B. (2015) Structure and function of Toll/interleukin-1 receptor/resistance protein (TIR) domains. *Apoptosis* **20**, 250–261.
3. Ve, T., Gay, N. J., Mansell, A., Kobe, B., and Kellie, S. (2012) Adaptors in toll-like receptor signaling and their potential as therapeutic targets. *Current Drug Targets* **13**, 1360–1374.
4. Kagan, J. C., and Medzhitov, R. (2006) Phosphoinositide-mediated adaptor recruitment controls Toll-like receptor signaling. *Cell* **125**, 943–955.
5. Valkov, E., Stamp, A., DiMaio, F., Baker, D., Verstak, B., Roversi, P., Kellie, S., Sweet, M. J., Mansell, A., Gay, N. J., Martin, J. L., and Kobe, B. (2011) Crystal structure of Toll-like receptor adaptor MAL/TIRAP reveals the molecular basis for signal transduction and disease protection. *Proceedings of the National Academy of Sciences of the U.S.A.* **108**, 14879–14884.
6. Lin, Z., Lu, J., Zhou, W., and Shen, Y. (2012) Structural insights into TIR domain specificity of the bridging adaptor Mal in TLR4 signaling. *PLoS One* **7**, e34202.
7. Snyder, G. A., Deredge, D., Waldhuber, A., Fresquez, T., Wilkins, D. Z., Smith, P. T., Durr, S., Cirl, C., Jiang, J., Jennings, W., Luchetti, T., Snyder, N., Sundberg, E. J., Wintrode, P., Miethke, T., and Xiao, T. S. (2014) Crystal structures of the Toll/Interleukin-1 receptor (TIR) domains from the *Brucella* protein TcpB and host adaptor TIRAP reveal mechanisms of molecular mimicry. *Journal of Biological Chemistry* **289**, 669–679.
8. Woo, J. R., Kim, S., Shoelson, S. E., and Park, S. (2012) X-ray Crystallographic structure of TIR-domain from the human TIR-domain containing adaptor protein/MyD88-adaptor-like protein (TIRAP/MAL). *Bulletin of the Korean Chemical Society* **33**, 3091–3094.
9. Dunne, A., Ejdeback, M., Ludidi, P. L., O'Neill, L. A., and Gay, N. J. (2003) Structural complementarity of Toll/interleukin-1 receptor domains in Toll-like receptors and the adaptors Mal and MyD88. *Journal of Biolocial Chemistry* **278**, 41443–41451.
10. Li, C., Zienkiewicz, J., and Hawiger, J. (2005) Interactive sites in the MyD88 Toll/interleukin (IL) 1 receptor domain responsible for coupling to the IL1beta signaling pathway. *Journal of Biological Chemistry* **280**, 26152–26159.
11. Jiang, Z., Georgel, P., Li, C., Choe, J., Crozat, K., Rutschmann, S., Du, X., Bigby, T., Mudd, S., Sovath, S., Wilson, I. A., Olson, A., and Beutler, B. (2006) Details of Toll-like receptor:adapter interaction revealed by germ-line mutagenesis. *Proceedings of the National Academy of Science U.S.A.* **103**, 10961–10966.

12. Valkov, E., Stamp, A., Dimaio, F., Baker, D., Verstak, B., Roversi, P., Kellie, S., Sweet, M. J., Mansell, A., Gay, N. J., Martin, J. L., and Kobe, B. (2011) Crystal structure of Toll-like receptor adaptor MAL/TIRAP reveals the molecular basis for signal transduction and disease protection. *Proceedings of the National Academy of Science U.S.A.* **108**, 14879–14884.
13. Eschenfeldt, W. H., Lucy, S., Millard, C. S., Joachimiak, A., and Mark, I. D. (2009) A family of LIC vectors for high-throughput cloning and purification of proteins. *Methods in Molecular Biology* **498**, 105–115.
14. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expression and Purification* **41**, 207–234.
15. Tyler, R. C., Sreenath, H. K., Singh, S., Aceti, D. J., Bingman, C. A., Markley, J. L., and Fox, B. G. (2005) Auto-induction medium for the production of [U-15N]- and [U-13C, U-15N]-labeled proteins for NMR screening and structure determination. *Protein Expression and Purification* **40**, 268–278.
16. Mobli, M., Stern, A. S., Bermel, W., King, G. F., and Hoch, J. C. (2010) A non-uniformly sampled 4D HCC(CO)NH-TOCSY experiment processed using maximum entropy for rapid protein sidechain assignment. *Journal of Magnetic Resonance* **204**, 160–164.
17. Mobli, M., Maciejewski, M. W., Gryk, M. R., and Hoch, J. C. (2007) An automated tool for maximum entropy reconstruction of biomolecular NMR spectra. *Nature Methods* **4**, 467–468.
18. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696.
19. Herrmann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *Journal of Molecular Biology* **319**, 209–227.
20. Guntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods in Molecular Biology* **278**, 353–378.
21. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *Journal of Biomolecular NMR* **13**, 289–302.
22. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of Biomolecular NMR* **44**, 213–223.

23. Gilbert, H. F. (1995) Thiol/disulfide exchange equilibria and disulfide bond stability. *Methods in Enzymology* **251**, 8–28.

24. Sharma, D., and Rajarathnam, K. (2000) 13C NMR chemcial shifts can predict disulfide bond formation. *Journal of Biomolecular NMR* **18**, 165–171.

25. Martin, O. A., Villegas, M. E., Vila, J. A., and Scheraga, H. A. (2010) Analysis of $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts of cysteine and cystine residues in proteins: a quantum chemical approach. *Journal of Biomolecular NMR* **46**, 217–225.

26. Schafer, F. Q., and Buettner, G. R. (2001) Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. *Free Radical Biology and Medicine* **30**, 1191–1212.

27. Toshchakov, V. U., Basu, S., Fenton, M. J., and Vogel, S. N. (2005) Differential involvement of BB loops of toll-IL-1 resistance (TIR) domain-containing adapter proteins in TLR4- versus TLR2-mediated signal transduction. *Journal of Immunology* **175**, 494–500.

28. Stack, J., and Bowie, A. G. (2012) Poxviral protein A46 antagonizes Toll-like receptor 4 signaling by targeting BB loop motifs in Toll-IL-1 receptor adaptor proteins to disrupt receptor:adaptor interactions. *Journal of Biological Chemistry* **287**, 22672–22682.

29. Waschutza, G., Li, V., Schafer, T., Schomburg, D., Villmann, C., Zakaria, H., and Otto, B. (1996) Engineered disulfide bonds in recombinant human interferon-gamma: the impact of the N-terminal helix A and the AB-loop on protein stability. *Protein Engineering, Design & Selection* **9**, 905–912.

30. Stewart, E. J., Aslund, F., and Beckwith, J. (1998) Disulfide bond formation in the *Escherichia coli* cytoplasm: an in vivo role reversal for the thioredoxins. *EMBO Journal* **17**, 5543–5550.

31. Li, Q., Harraz, M. M., Zhou, W., Zhang, L. N., Ding, W., Zhang, Y., Eggleston, T., Yeaman, C., Banfi, B., and Engelhardt, J. F. (2006) Nox2 and Rac1 regulate H(2)O(2)-dependent recruitment of TRAF6 to endosomal interleukin-1 receptor complexes. *Molecular and Cellular Biology* **26**, 140–154.

32. Ngkelo, A., Meja, K., Yeadon, M., Adcock, I., and Kirkham, P. A. (2012) LPS induced inflammatory responses in human peripheral blood mononuclear cells is mediated through NOX4 and G(i)α dependent PI-3kinase signalling. *Journal of Inflammation (London, England)* **9**, 1–1.

33. Reynaert, N. L., van der Vliet, A., Guala, A. S., McGovern, T., Hristova, M., Pantano, C., Heintz, N. H., Heim, J., Ho, Y. S., Matthews, D. E., Wouters, E. F., and Janssen-Heininger, Y. M. (2006) Dynamic redox control of NF-kappaB through glutaredoxin-regulated *S*-

glutathionylation of inhibitory kappaB kinase beta. *Proceedings of the National Academy of Science U.S.A* **103**, 13086–13091.

34. Menon, D., and Board, P. G. (2013) A Role for glutathione transferase omega 1 (GSTO1-1) in the glutathionylation cycle. *Journal of Biological Chemistry* **288**, 25769–25779.

35. Menon, D., Coll, R., O'Neill, L. A. J., and Board, P. G. (2014) Glutathione transferase Omega 1 is required for the lipopolysaccharide-stimulated induction of NADPH oxidase 1 and the production of reactive oxygen species in macrophages. *Free Radical Biology and Medicine* **73**, 318–327.

36. Menon, D., Coll, R., O'Neill, L. A. J., and Board, P. G. (2015) GSTO1-1 modulates metabolism in macrophages activated through the LPS and TLR4 pathway. *Journal of Cell Science* **128**, 1982–1990.

37. Nagpal, K., Plantinga, T. S., Wong, J., Monks, B. G., Gay, N. J., Netea, M. G., Fitzgerald, K. A., and Golenbock, D. T. (2009) A TIR domain variant of MyD88 adapter-like (Mal)/TIRAP results in loss of MyD88 binding and reduced TLR2/TLR4 signaling. *Journal of Biological Chemistry* **284**, 25742–25748.

38. Vyncke, L., Bovijn, C., Pauwels, E., Van Acker, T., Ruyssinck, E., Burg, E., Tavernier, J., and Peelman, F. (2016) Reconstructing the TIR side of the myddosome: a paradigm for TIR-TIR interactions. *Structure* **24**, 437–447.

39. Paracha, R. Z., Ali, A., Ahmad, J., Hussain, R., Niazi, U., and Muhammad, S. A. (2014) Structural evaluation of BTK and PKCdelta mediated phosphorylation of MAL at positions Tyr86 and Tyr106. *Computational Biology and Chemistry* **51**, 22–35.

40. Guven-Maiorov, E., Keskin, O., Gursoy, A., VanWaes, C., Chen, Z., Tsai, C. J., and Nussinov, R. (2015) The architecture of the TIR domain signalosome in the Toll-like receptor-4 signaling pathway. *Scientific Reports* **5**, 13128.

Chapter 4

Solution structure of Sr33 using NMR spectroscopy

4.1 Contribution

Chapter 4 includes the manuscript titled "*The solution structure of Sr33 challenges paradigms for coiled-coil domain dimerization in plant NMR immune receptors*," which will be submitted imminently for publication. This piece of work has involved a large collaboration between researchers with skills spanning structural biology, biophysics and plant biology techniques. My role in this work includes:

- The expression and purification of the ^{15}N and ^{13}C labelled constructs and preparation of this protein for NMR spectroscopy
- Assignment of ^{1}H, ^{15}N and ^{13}C resonances to the atoms of $Sr33^{6-120}$
- Refinement of the $Sr33^{6-120}$ structure
- Writing of the NMR sections of the paper and figure preparation
- Drafting of the entire paper in combination with other authors of the manuscript

The NMR structure of $Sr33^{6-120}$ was a very important outcome for the full biophysical, structural and functional studies. This is recognised by my joint-first author position on the manuscript.

4.2 Preface

One research focus in the Kobe laboratory is the structural characterisation of plant NLR (nucleotide-binding oligomerization domain (NOD)-like receptor) proteins. Plant NLRs are important innate immunity receptors capable of detecting molecules from plant pathogens, which the pathogen utilises during invasion, to initiate defence pathways. During my PhD candidature, an opportunity arose to be a part of a large collaboration that would focus on understanding the signalling function of the N-terminal coiled-coil (CC) domain from the agriculturally important wheat stem-rust resistance NLR, Sr33. Plant NLR proteins generally contain either a CC or TIR domain at their N-terminus and both have been implicated in defence signalling through homotypic interactions. Currently in the literature, there are two published crystal structures of CC domains from NLR proteins, namely: barley MLA10, which is responsible for resistance to barley powdery mildew; and potato Rx, responsible for resistance to potato virus X. The crystal structures of the CC domains from the two proteins are significantly different. The structure of MLA10 revealed an antiparallel homodimer adopting a helix-loop-helix fold, while the structure of Rx revealed a compact monomeric four-helical bundle. This data suggested that the CC domains of plant NLRs are

structurally diverse, which could have implications for the function of this domain in defence signalling.

Sr33 is an ortholog of MLA10 and the CC domains share a high sequence identity (~82%). Both Sr33 and MLA10 only share limited similarity with the distantly related Rx (~18%). Multi-angle laser light scattering (MALS) experiments performed in our laboratory indicated that the CC domains from all three proteins, MLA10, Sr33 and Rx, were monomeric in solution, which was not consistent with the reported crystal structure and biochemical data for MLA10 (12). This initiated a structural and biophysical study of Sr33, which included x-ray crystallography, small-angled x-ray scattering and NMR, where I made my key contribution. I determined the NMR structure of Sr33 CC domain, which revealed that the structure of Sr33 CC domain has a fold similar to that of Rx, adopting a compact four-helix bundle. This is strikingly different to the structure of the more closely related MLA10. In addition, this NMR study, with support from other biophysical studies detailed in the manuscript below, has shown that Sr33 is monomeric and has little propensity to self-associate *in vitro*. This work has then led to the determination of a minimal functional unit of an NLR coiled-coil domain, which is ~20 amino acids longer than the previously presented structures and the Sr33 structure presented here. This has revealed that *in planta*, these longer CC domains self-associate and act as a switch that result in a defence activation, represented by an apoptotic-like cell death response. The structural discovery of Sr33 presented here suggests that distant and closely related CC domains are structurally conserved and have reconciled previous conflicting structural data.

The solution structure of Sr33 challenges paradigms for coiled-coil domain dimerization in plant NLR immune receptors

Lachlan W. Casey[1]*, **Peter Lavrencic**[1,5]*, Adam Bentham[1,2]*, Stella Cesari[3], Daniel Ericsson[4], Peter A. Anderson[2], Alan E. Mark[1], Peter N. Dodds[3], Mehdi Mobli[5§], Bostjan Kobe[1§] and Simon J. Williams[1,2,6§]

[1]School of Chemistry and Molecular Biosciences, Institute for Molecular Bioscience and Australian Infectious Diseases Research Centre, University of Queensland, Brisbane, Queensland 4072, Australia. [2]School of Biological Sciences, Flinders University, Adelaide, SA 5001, Australia. [3]CSIRO Agriculture Flagship, GPO Box 1600, Canberra ACT 2601, Australia [4]MX Beamlines, Australian Synchrotron, 800 Blackburn Road, Clayton, Victoria 3168, Australia. [5]Centre for Advanced Imaging, University of Queensland, Brisbane, Queensland 4072, Australia. [6]Plant Sciences Division, Research School of Biology, The Australian National University, Canberra 2601, Australia

* These authors contributed equally to this work

[§]***Corresponding authors:*** s.williams8@uq.edu.au (+61733469073), b.kobe@uq.edu.au (+61733652132), m.mobli@uq.edu.au (+617334 60352)

Classification: BIOLOGICAL SCIENCES, Biophysics

Key words: Plant innate immunity, resistance (R) protein, coiled coil (CC) domain, nucleotide-binding oligomerization domain (NOD)-like receptor (NLR), nuclear magnetic resonance (NMR), small-angle X-ray scattering (SAXS), multi-angle laser light scattering (MALS), X-ray crystallography.

4.3 Abstract

Plant NLRs (nucleotide-binding oligomerization domain (NOD)-like receptors) recognize specific pathogen effector proteins and initiate immune responses. Their multi-domain architecture comprises a central nucleotide-binding domain, a C-terminal leucine-rich repeat domain and either CC (coiled-coil) or TIR (Toll/interleukin-1 receptor/resistance) domain at the N-terminus. The CC and TIR domains are involved in defence signalling and have been implicated in homotypic interactions. Here, we present the solution structure of the coiled-coil fragment from the wheat stem-rust NLR protein Sr33 comprising residues 6–120. The domain adopts a four-helix bundle conformation that unexpectedly differs significantly from the published crystal structure of the CC domain from the

orthologous barley powdery mildew resistance protein, MLA10, which forms a homo-dimer. However, the structure of Sr33 is similar to the distantly related CC domain from the potato NLR protein Rx. Biophysical studies demonstrate that the analogous regions of Sr33, MLA10 and Rx are monomeric and adopt similar folds in solution. However, we found that larger N-terminal fragments of Sr33 and MLA10 (up to 160 amino acids) can homo-dimerise in solution and co-immunoprecipitate when expressed *in planta*, and this self-association property correlates with their ability to signal cell death *in planta*. We define the minimal functional signalling region of the CC domain in these proteins as comprising amino acids 1–142. This includes an additional 22 amino acids predicted to form an α-helix that is absent from the crystal and solution structures and is required for both dimerization and signalling. The correlation between signalling and self-association suggest that collectively, our study contributes new insights into our structural and functional understanding of the CC domains from plant NLR proteins.

4.4 Introduction

Plant diseases constitute a major economic and social burden worldwide, and the appearance of new or resistant pathogens can pose significant challenges. Plants rely on their innate immune systems to combat emerging pathogens. An important component of plant innate immunity is the recognition of pathogen effector molecules by resistance (R) proteins within the plant cell. R protein activation triggers a process known as the hypersensitive response (HR), which often culminates in localized cell death at the site of infection, leading to general immunity of the whole plant (1,2).

One such resistance protein is encoded by the recently discovered wheat gene *Sr33* (3). *Sr33* confers resistance to the virulent Ug99 strain of wheat stem rust, *Puccinia graminis f. sp. tritici* (*Pgt*), a pathogen well recognized for its potential threat to global food security. As a member of the canonical class of plant R genes, *Sr33* encodes a large (~110 kDa) multi-domain protein consisting of a central nucleotide-binding (NB) domain, a C-terminal leucine-rich repeat (LRR) domain and an N-terminal coiled-coil (CC) domain (3). These proteins have a similar domain arrangement and function to the nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs) from mammals, and are commonly referred to as plant NLRs.

While the molecular detail of plant NLR activation and signalling is not fully understood, targeted studies have helped define the roles of their different domains. The central NB domain appears to control the activation of the protein through nucleotide binding and exchange (4-6). The LRR domain

plays a role in effector recognition specificity for a number of plant NLRs, and in some cases it is implicated in effector binding (7,8). The LRR domain also appears to have a general autoinhibitory role and structural and biochemical studies of the human NLR proteins NAIP and NLRC4 support this conclusion (9). In plant NLRs, the N-terminal domain consists of either a CC (coiled-coil) domain, as in *Sr33*, or a TIR (Toll/interleukin-1 receptor/resistance protein) domain. Both CC and TIR domains are implicated in downstream signalling, and have been shown to be necessary and sufficient for HR responses in a number of systems (10-16).

Oligomerization has long been implicated in animal NLR protein activation and function and this process has been recently observed structurally for the activated inflammasome formed by the animal NLR proteins, NAIP2/NLRC4 (17-19). Our understanding of these processes in plant NLRs is much more limited. To date, effector-induced self-association of a full-length plant NLR has been demonstrated only for the tobacco mosaic virus resistance protein, N (20). Structure-guided studies restricted to the N-terminal TIR or CC domains have expanded upon this, demonstrating that weakly associating, transient self-interactions are necessary for signalling by the TIR domains of the plant NLRs L6 and RPS4 (11,15). Furthermore, strong heterodimeric interactions between the TIR domains of RPS4/RRS1 appear to negatively regulate this dual-resistance protein complex (15).

Even less is known about oligomerization of the CC domain-containing plant NLRs like Sr33. While the plant NLR TIR domains have a conserved fold (11,15), the structures of two (partial) CC domains are strikingly different. The N-terminal amino-acids 5–120 of barley powdery mildew A 10 (MLA10) crystallized as an antiparallel homodimer adopting a helix-loop-helix fold (12). The structure appears to form an obligate dimer. This suggested that dimerization is required for activation, although this has not been established experimentally. The second structure, the CC domain of potato Rx, was seen to adopt a compact four-helical bundle in a 1:1 hetero-association with the RanGAP2 co-factor (21), clearly distinct from the MLA10 homodimer. Moreover, the Rx CC was stable as a monomer without its cofactor *in vitro*. Together, these studies suggested that significant structural and mechanistic variation may be present among the CC domains of plant NLRs.

To understand the role of CC domains in NLR protein signalling further, we investigated the wheat stem-rust NLR protein Sr33. Here, we present the three-dimensional (3D) solution structure of Sr33 CC domain (residues 6–120; $Sr33^{6\text{-}120}$), determined by nuclear magnetic resonance (NMR) spectroscopy. Despite being an ortholog of MLA10 and having high sequence identity (~82%) across this region, our structure looks much more similar to the CC domain (residues 1–122) of Rx ($Rx^{1\text{-}122}$; ~18% identity). Subsequently, we carried out a detailed biophysical comparison of the CC domains of Sr33, MLA10 and Rx. These studies suggest that the CC-NLRs may not be as disparate as previously

thought. We show that the regions responsible for self-association are beyond the domain boundaries set by current structural data and in doing so, we define the minimal functional region of plant CC domains.

4.5 Methods

4.5.1 *Cloning, expression and purification*

The cDNAs coding for the proteins under study were cloned into the pMCSG7 vector by ligation-independent cloning (22). For biophysical studies, the proteins were expressed in *Escherichia coli* BL21 (DE3) using the autoinduction method (23). Cells were lysed via sonication in the lysis buffer (consisting of 50 mM HEPES pH 8.0, 300 mM NaCl and 1 mM dithiothreitol (DTT)) for $Sr33^{6-120}$, $MLA10^{5-120}$ and Rx^{1-122}. A similar lysis buffer was used for the longer CC domain fragments (corresponding to $Sr33^{6-144}$, $Sr33^{6-160}$, $MLA10^{5-144}$ and $MLA10^{5-160}$); however, the pH was adjusted to 7.5 and 500 mM of NaCl was used. The proteins were separated from clarified cell lysate via immobilized metal affinity chromatography (IMAC), facilitated by N-terminal 6 x histidine tags. Overnight treatment with TEV (tobacco etch virus) protease was used to remove the histidine tag, leaving a three-residue N-terminal overhang (Ser-Asp-Ala). The theoretical molecular weights for $Sr33^{6-120}$, $MLA10^{5-120}$ and Rx^{1-122} are therefore 13.1, 13.4 and 14.2 kDa, respectively. The cleaved protein was reapplied to the nickel affinity chromatography column to remove the TEV protease and other contaminants. The proteins were further purified using a Superdex 75 HiLoad 26/60 size-exclusion chromatography column [GE Healthcare] equilibrated with 10 mM HEPES pH 8.0, 150 mM NaCl and 1 mM DTT.

4.5.2 *Protein expression for NMR spectroscopy*

E. coli BL21 cells expressing the $Sr33^{6-120}$ protein (see above) were grown in M9 minimal media containing ^{13}C-labelled glucose, and ^{15}N-labelled ammonium chloride. Protein expression was induced using 1 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) at 20˚C for overnight protein expression. The $^{13}C/^{15}N$-labelled $Sr33^{6-120}$ protein was purified using nickel affinity and size-exclusion chromatography as described above.

The correlation time of the protein was estimated based on transverse relaxation rates (T_2), measured as described previously (24). The correlation time was converted to a molecular weight

using the Stoke-Einstein equations (25), using a modified equation for estimation of protein volumes (26) with the addition of 2 Å to account for the hydration shell.

4.5.3 *NMR data acquisition*

The $^{13}C/^{15}N$-labelled $Sr33^{6-120}$ sample containing 5% D_2O was filtered using a low-protein-binding Ultrafree-MC centrifugal filter (0.22 μm pore size; Millipore, MA, USA), then 300 μL was added to a susceptibility-matched 5 mm outer-diameter microtube [Shigemi Inc., Japan].

NMR data were acquired at 25˚C using a 900 MHz AVANCE spectrometer [Bruker BioSpin, Germany] equipped with a cryogenically cooled probe. Data used for resonance assignment were acquired using non-uniform sampling (NUS); sampling schedules that approximated the rate of signal decay along the various indirect dimensions were generated using sched3D (27). The decay rates used were 1 Hz for all constant-time ^{15}N dimensions, 30 Hz for all ^{13}C dimensions, and 15 Hz for the semi-constant indirect ^{1}H dimension. ^{13}C- and ^{15}N-edited HSQC-NOESY experiments were acquired using linear sampling. Separate experiments were acquired for the aliphatic and aromatic regions of the ^{13}C dimension.

NUS data were processed using the Rowland NMR toolkit (www.rowland.org/rnmrtk/toolkit.html); maximum entropy parameters were selected automatically as described previously (28,29). NMR spectra were analyzed and assigned using the program CcpNmr (30). $^{1}H_N$, ^{15}N, ^{13}C backbone resonance assignments were obtained from the analysis of amide-proton strips in 3D HNCACB, CBCA(CO)NH, and HNCO spectra. Sidechain ^{1}H and ^{13}C chemical shifts were obtained primarily from 3D H(CC)(CO)NH-TOCSY and (H)CC(CO)NH-TOCSY spectra, respectively. The remaining side-chain assignments were derived from 3D H(C)CH-TOCSY and ^{15}N- and ^{13}C-edited NOESY-HSQC spectra.

4.5.4 *NMR structure determination*

Distance restraints for structure calculations were derived from 3D ^{13}C- and ^{15}N-edited NOESY-HSQC spectra acquired with a mixing time of 120 ms. NOESY spectra were manually peak-picked and integrated using the box-sum method in CcpNMR (31). The peak lists were then assigned and an ensemble of structures calculated automatically using the torsion angle dynamics package CYANA (31,32). The tolerances used in the structure calculations were 0.03 ppm in the indirect ^{1}H

dimension, 0.02 ppm in the direct ^{1}H dimension, 0.2 ppm for the aromatic ^{13}C and ^{15}N dimensions, and 0.4 ppm for the aliphatic ^{13}C data.

Backbone dihedral-angle restraints (112 restraints each for ϕ and ψ) were derived from TALOS+ chemical shift analysis (33); the restraint range was set to twice the estimated standard deviation. All X-Pro peptide bonds were clearly identified as *trans* on the basis of characteristic C_β and C_γ chemical shifts for the Pro residues.

CYANA was used to calculate 200 structures from random starting conformations, then the 20 conformers with the lowest CYANA target function were chosen to represent the structural ensemble. During the automated NOESY assignment/structure calculation process CYANA assigned 94.4% of all NOESY crosspeaks (3186 out of 3372) for Sr33.

4.5.5 *Size-exclusion chromatography (SEC)-coupled multi-angle light scattering (MALS)*

SEC-MALS was performed using an in-line Superdex 200 100/300 GL or Superdex 200 Increase 5/150 GL SEC column [GE Healthcare] combined with a Dawn Heleos II 18-angle light-scattering detector coupled with an Optilab TrEX refractive index detector [Wyatt Technology, Santa Barbara, CA, USA]. Purified proteins were separated at 0.5 mL/min (10/300) or 0.25 mL/min (5/150) in 10 mM HEPES pH 8.0 and 150 mM NaCl. Molecular-mass calculations were performed using the Astra6.1 software [Wyatt Technology]. Input of the refractive increment (dn/dc values) was set at 0.186 in the molecular-mass calculations, based on the premise that dn/dc is constant for unmodified proteins (52). The molecular mass was determined across the protein elution peak.

4.5.6 *Analytical size-exclusion chromatography (SEC) and cross-linking*

The purified MLA10$^{5\text{-}120}$ protein (450 μg) was separated on a Superdex 75 10/300 GL SEC column with a mobile phase consisting of 10 mM HEPES pH 7.5 and 150 mM NaCl (SEC buffer). Protein size markers chymotrypsin (25 kDa) and cytochrome c (15 kDa) were separated using the same conditions as for MLA10$^{5\text{-}120}$. Cross-linking experiments were performed as previously described (12). In brief, 20 μL of MLA10$^{5\text{-}120}$ (in the SEC buffer) at a concentration of 150 μM was mixed with 5 μL of BS3 (bis(sulfosuccinimidyl)suberate) at a concentration of 20 mM. The reaction was incubated on ice and monitored at various time points from 0-120 minutes. The reaction was quenched with equal volumes of 1 M Tris pH 7.5, before the samples were separated using 13% SDS-PAGE.

4.5.7 *Size-exclusion chromatography (SEC)-coupled small-angle X-ray scattering (SAXS)*

SEC-SAXS was performed during two shifts at the SAXS/WAXS beamline of the Australian Synchrotron on a Pilatus 1M detector, using an in-line WTC-030S5 SEC column and a 2 mL WTC-030S5G pre-column [Wyatt Technology], together with a Prominence modular HPLC system [Shimadzu Scientific Instruments]. All experiments were conducted at 16°C using 10 mM HEPES (pH 7.5), 150 mM NaCl buffer with 1 mM DTT. For data collection, eluate from the column was directed through a 1 mm quartz capillary mounted in the beam. For all samples, the injected volume was 95 µL at 30 mg/mL protein concentration, as determined by UV absorbance at 280 nm. High concentrations were used to maximize signal after dilution during gel-filtration.

The data for Sr33 was collected in 5 s exposures at 0.05 s intervals with a flow rate of 0.2 mL/min. A Wyatt WTC-030S5G pre-column was used upstream of the WTC-030S5. The sample-to-detector distance was 1.6 m, and a wavelength of 1.12713 Å yielded a range of momentum transfer ($0.009 < q < 0.478$ Å^{-1}, where $q = 4\pi.\sin(\vartheta)/\lambda$). The data for MLA10$^{5\text{-}120}$ and Rx$^{1\text{-}122}$ were collected on the same column at a different time, in 2 s exposures at 0.05 s intervals, with sample flowing at a rate of 0.5 mL/min. The pre-column was not used for these samples. In this case, a sample-to-detector distance of 1.4 m was used to obtain data over the range $0.010 < q < 0.614$ Å^{-1}. Data reduction and subtraction was performed using scatterBrain (http://www.synchrotron.org.au/index.php/aussyncbeamlines/saxswaxs/software-saxswaxs). Unless noted otherwise, subsequent analyses were performed using the tools in version 2.6 of the ATSAS program suite (34).

100 frames immediately preceding each peak were summed and normalized for exposure time to obtain buffer blanks. Initially, these buffers were subtracted from each individual image to generate a series of subtracted frames across the elution peak, in order to evaluate the evolution of calculated properties. To this end, $I(0)$ and R_g were calculated for each frame using the Guinier approximation, as implemented in batch-mode AUTORG, for points such that $q.R_g < 1.3$. Molecular weights were calculated using the volume of correlation (V_c) (35).

These metrics were evaluated for variation across the peak. To obtain the final scattering curves for analysis, the original images, from peak regions over which R_g and MW_{Vc} remained stable, were summed and normalized in scatterBrain, and then subtracted from the corresponding blank.

Guinier analysis and the determination of $I(0)$, R_g and MW_{Vc} were performed on the summed and averaged curves in the same manner as for individual frames, but using AUTORG in Primus instead of in batch (36). Data-points closer to the beamstop than the first selected Guinier point were

discarded. Points where q > 0.46 $Å^{-1}$ were also discarded, due to poor signal-to-noise. Distance distributions, *P(r)*, were then obtained by indirect transformation in GNOM (37), informed by AUTOGNOM. In addition to MW_{Vc}, molecular weights were also estimated from the Porod volume calculated by GNOM. In this estimation, the protein density was assumed to take the canonical value of 1.37 g cm^{-1} (38).

Theoretical scattering was calculated from atomic models using FoXS (39). Short stretches of residues not visible in the electron density of the published $MLA10^{5-120}$ crystal structure were added to both chains using the loop-building routines in MODELLER (40,41) independently from the SAXS data.

4.5.8 *Crystallization and crystal structure determination of $MLA10^{5-120}$*

Native and selenomethionine-labelled $MLA10^{5-120}$ protein at 10 mg/mL and 6 mg/mL, respectively, in 10 mM HEPES (pH 8.0), 100 mM NaCl, and 1 mM DTT were used in crystallization trials. Crystallization experiments were initially performed with native protein using hanging-drop vapour diffusion in 96-well plates. Several commercial screens were used, including Index, PEG/Ion and PEGRx [Hampton Research] and Pact Premier and JCSG+ [Molecular Dimensions]. 100 nl protein solution and 100 nl well solution were prepared on hanging-drop seals [TTP4150-5100 sourced from Millennium Science, Australia] using a Mosquito robot [TTP Lab- Tech, UK] and equilibrated against 75 ml reservoir solution. The drops were monitored and imaged using the Rock Imager system [Formulatrix, USA]. Numerous promising hits were observed within 24 hours; however, the crystals grown in Pact Premier, condition B4 (MIB buffer pH 7.0 (42), 25% PEG 1500) were pursued for data collection. Crystals grown in larger 1:1 μL (protein: well solution) drops were cryo-protected using the well-solution containing 20% glycerol prior to flash-cooling in liquid nitrogen. X-ray diffraction data of native crystals were collected from a single crystal at the Australian Synchrotron MX1 beamline to ~2.0 Å resolution using a wavelength of 0.9537 Å. The crystal-to-detector distance was set to 200 mm and the oscillation range was 0.5°. Data collection was performed using Blu-Ice software (43), indexed and integrated using XDS (44) and scaled with AIMLESS within the CCP4 suite (45). With the native dataset molecular replacement was attempted using the published $MLA10^{5-120}$ structure (PDB ID 3QFL;(12)) in monomeric, dimeric and various truncated forms, as well as the structure of Rx^{1-122} (PDB ID 4M70; (21)); however, a solution could not be obtained. Subsequently, selenomethionine-labelled protein (confirmed by mass spectrometry, data not shown) was crystallized as described for the native protein. X-ray diffraction data of selenomethionine-labelled crystals were collected from a single crystal at the Australian Synchrotron MX2 beamline to ~2.1 Å

resolution using a wavelength 0.9792 Å. The crystal-to-detector distance was set to 200 mm and the oscillation range was 0.5°. Data collection was performed using Blu-Ice software (43), indexed and integrated using XDS (44), and scaled with AIMLESS within the CCP4 suite (45).

The crystals of MLA10$^{5\text{-}120}$ appeared to have symmetry of space group P2 21 21 and the structure was solved using single-wavelength anomalous diffraction (SAD) through the CRANK2 pipeline (46). Due to issues during refinement in the P2 21 21 space group, we performed further refinement in P1. Statistics for the refined atomic model are presented in Appendix Table 4.3.

4.5.9 *Constructs for in planta analyses*

For transient expression in *N. benthamiana*, molecular cloning was performed by a combination of Quikchange site-directed mutagenesis [Agilent Technologies] and Gateway recombination [Life Technologies]. The MLA10$^{1\text{-}160}$, Sr33$^{1\text{-}160}$ and Sr50$^{1\text{-}163}$ constructs cloned in pDONR207 (29) were used as templates for site-directed deletion to generate the MLA10$^{1\text{-}130}$, MLA10$^{1\text{-}135}$, MLA10$^{1\text{-}141}$, MLA10$^{1\text{-}142}$, MLA10$^{1\text{-}144}$, MLA10$^{1\text{-}148}$, Sr33$^{1\text{-}130}$, Sr33$^{1\text{-}135}$, Sr33$^{1\text{-}141}$, Sr33$^{1\text{-}142}$, Sr33$_{1\text{-}144}$, Sr33$_{1\text{-}148}$, Sr50$_{1\text{-}133}$, Sr50$^{1\text{-}138}$, Sr50$^{1\text{-}144}$, Sr50$^{1\text{-}145}$, Sr50$^{1\text{-}147}$ and Sr50$^{1\text{-}151}$ ENTRY constructs. These constructs were then recombined by LR reaction in the binary vector pBIN19-35S::GTW:3HA or pBIN19-35S::GTW:CFP by LR coning to obtain expression vectors.

4.5.10 *Transient protein expression and cell death assays in N. benthamiana*

N. benthamiana plants were grown in a growth chamber at 23°C with a 16 hours light period. For *N. benthamiana* leaf transformations, pBIN19-derived vector constructs were transformed into *Agrobacterium tumefaciens* strain GV3101_pMP90. Bacterial strains were grown in Luria-Bertani liquid medium containing 50 mg/ml rifampicin, 15 mg/ml gentamycin and 25 mg/ml kanamycin at 28°C for 24 hours. Bacteria were harvested by centrifugation, resuspended in infiltration medium (10 mM MES pH 5.6, 10 mM MgCl2 and 150 µM acetosyringone) to an OD600 nm ranging from 0.5 to 1, and incubated for 2 hours at room temperature before leaf infiltration. Three leaves from two plants were infiltrated for each combination of constructs and the experiment was repeated three times independently. The infiltrated plants were incubated in growth chambers under controlled conditions for all following assays. For documentation of cell death, leaves were scanned five days after infiltration.

4.5.11 *Protein extraction Western blot and co-immunoprecipitation*

Protein extraction from *N. benthamiana* leaves and co-IP experiments were performed as described (49). For immunoblotting analysis, proteins were separated by SDS-PAGE and transferred to a nitrocellulose membrane. Membranes were blocked in 5% skimmed milk and probed with anti-HA-HRP antibodies [Roche] or anti-GFP antibodies [Roche] followed by goat anti-mouse antibodies conjugated with horseradish peroxidase [Pierce]. Labeling was detected using the SuperSignal West Femto chemiluminescence kit [Pierce]. Membranes were stained with Ponceau S to confirm equal loading.

4.6 Results

4.6.1 *The NMR structure of Sr33$^{6\text{-}120}$ reveals a compact four-helical bundle*

Previous structural studies of the Sr33 ortholog MLA10 encompassed residues 5–120 of the CC domain (MLA10$^{5\text{-}120}$). For our investigation, we initially focussed on an equivalent region within the Sr33 CC domain. Soluble Sr33 CC domain protein was produced by recombinant expression of a construct comprising residues 6–120 (hereafter referred to as Sr33$^{6\text{-}120}$) in *Escherichia coli*. The *E. coli* cells were grown in M9 minimal media and labelled selectively with ^{15}N and ^{13}C. The labelled protein was purified to homogeneity using metal-affinity chromatography, followed by size-exclusion chromatography (SEC). The atomic structure of Sr33$^{6\text{-}120}$ was determined by NMR spectroscopy.

Sr33$^{6\text{-}120}$ is monomeric in the buffer used for structural studies, according to average T_2 relaxation rates of the backbone amides (yielding the correlation time of ~8.7 ns, consistent with a calculated molecular weight ~13 kDa; the theoretical molecular weight of Sr33$^{6\text{-}120}$ is 13.1 kDa (24)). This is also evident in the sharp line-widths of the resonances, which are consistent with a monomeric and not a dimeric form of Sr33$^{6\text{-}120}$.

Analysis of the assigned chemical shifts (47) revealed four distinctive α-helical regions (α1, residues 7–19; α2, 28–51; α3, 60–87; and α4; 99–115). Other than the termini, two small regions (comprising residues 22–23 and 88–91) were found to have near random-coil chemical shifts, indicating that these regions are highly dynamic. The residues corresponding to 88–91 in MLA10$^{5\text{-}120}$ are poorly defined in the crystal structure as well; however, the residues corresponding to 22–23 in MLA10$^{5\text{-}120}$ appear in an ordered helical region (see below).

Dihedral angles derived from the chemical shift analysis were used together with distance restraints from ^{15}N and ^{13}C edited NOESY experiments to calculate a 3D structure of the protein. The structure shows that the protein is predominantly α-helical, folded in a four-helix bundle (Figures 4.1A and S4.2). The 20 lowest energy structures show an RMSD (root-mean-square-distance) for the amide backbone atoms (N, C_α, and C') of residues 6–89, 98–110 of 0.93 Å (Appendix Table 4.1).

Previous studies have indicated that the EDVID motif is conserved among CC domains and is important in mediating intramolecular interactions (21,48). In $Sr33^{6-120}$, the equivalent motif (residues 77–81) encodes residues EDAVD, which reside in the α3 helix with residues E77, D78, V80 and D81 all surface-exposed (Figure 4.1A). Despite the higher sequence similarity between $Sr33^{6-120}$ and $MLA10^{5-120}$, the solution structure of $Sr33^{6-120}$ resembles the structure of Rx^{1-122} more closely (Figure 4.1 B–E).

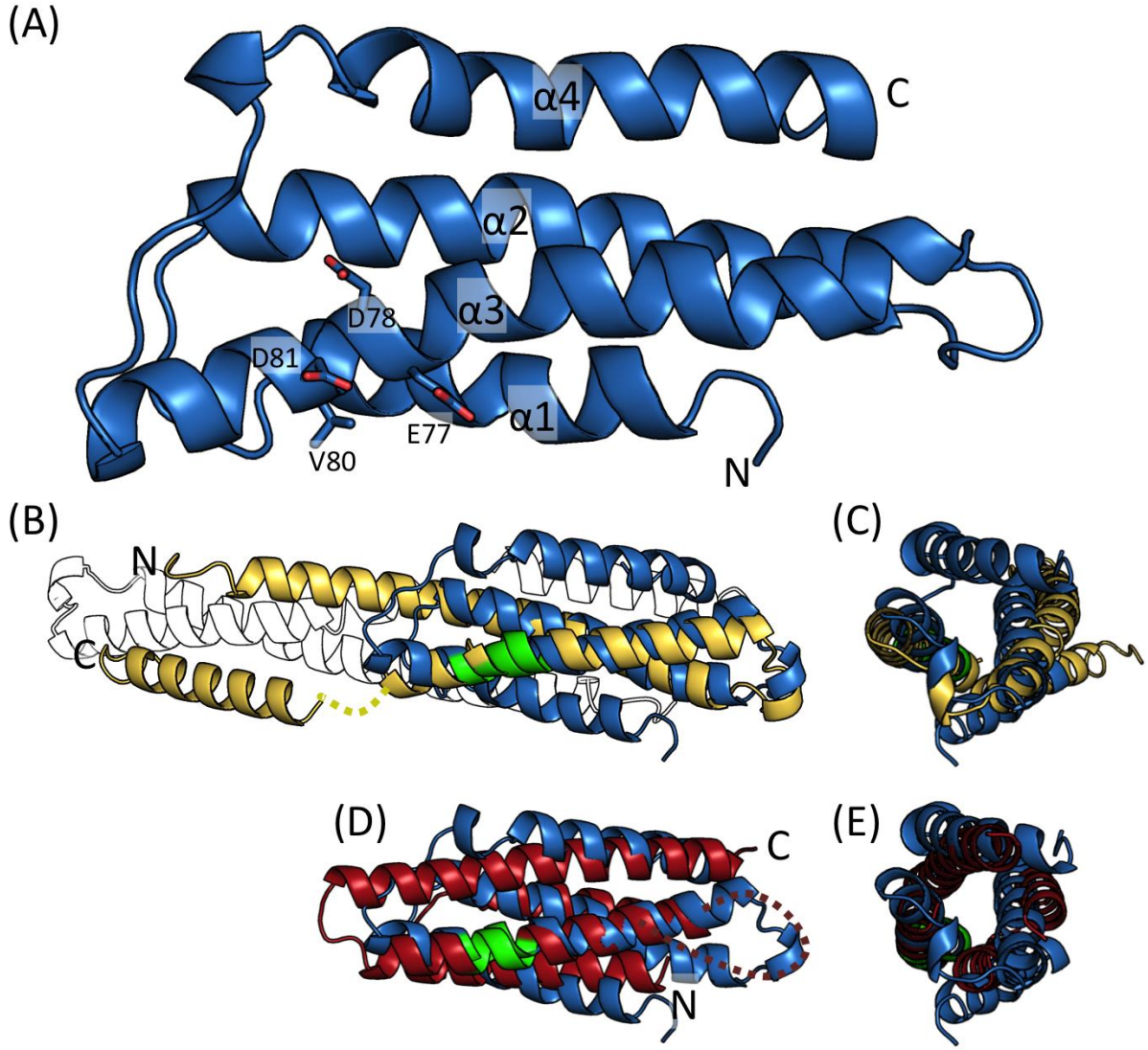


Figure 4.1 The solution structure of Sr33 reveals a four-helix bundle fold. (**A**) NMR structure of $Sr33^{6-120}$ in cartoon representation, with the individual helices and N- and C-termini labelled. The conserved EDVID motif (EDAVD in Sr33), is shown in stick representation. (**B**) Superposition of the $Sr33^{6-120}$ structure (blue) and the crystal structure of $MLA10^{5-120}$ (yellow) in cartoon representation. Missing residues in $MLA10^{5-120}$ structure (amino acids 91–95) are shown by a dotted line. Green colouring represents the position of the EDVID motif. The crystallographic dimer observed for $MLA10^{5-120}$ is shown as a black-and-white outline. (**C**) Superposition, as shown in B, rotated 90° around the y-axis. (**D**) Superposition of the $Sr33^{6-120}$ structure (blue) and the crystal structure of Rx^{1-122} (red) in cartoon representation. Missing residues in Rx^{1-122} structure (amino acids 40–50) are shown by a dotted line. Green colouring represents the position of the conserved EDVID motif.

4.6.2 *Biophysical characterization shows that Sr33$^{6\text{-}120}$, MLA10$^{5\text{-}120}$ and Rx$^{1\text{-}122}$ are predominantly monomeric in solution*

In light of the NMR results suggesting that Sr33$^{6\text{-}120}$ behaves as a monomer, we employed additional *in vitro* biophysical techniques to interrogate this further. We also produced and analyzed the MLA10$^{5\text{-}120}$ and Rx$^{1\text{-}122}$ constructs used for structure determination previously (12,21). SEC-coupled multi-angle light scattering (SEC-MALS) experiments performed on the three CC-domain proteins confirmed that Sr33$^{6\text{-}120}$ and Rx$^{1\text{-}122}$ were monomeric in solution, even at loading concentrations of 30 mg/mL (Figures 4.2A–C). Unexpectedly, this was also true of MLA10$^{5\text{-}120}$. The average MALS-derived molecular weights were 13.7 kDa for Sr33$^{6\text{-}120}$, 13.3 kDa for MLA10$^{5\text{-}120}$, and 13.3 kDa for Rx$^{1\text{-}122}$, while the predicted monomeric molecular weights from these constructs are 13.1 kDa, 13.4 kDa and 14.3 kDa, respectively.

We examined the oligomeric states further using in-line SEC-coupled small-angle X-ray scattering (SEC-SAXS), which revealed some variation that was not apparent by SEC-MALS (Figures 4.2D–F). The SAXS-derived molecular weight of Sr33$^{6\text{-}120}$ varied between 12.5 and 13.8 kDa across the peak, with a trend towards smaller particles at the elution tail. This variation was more pronounced for both MLA10$^{5\text{-}120}$ and Rx$^{1\text{-}122}$, with the calculated molecular weights increasing from ~13 at the start of the peak to ~16 kDa at the peak centre, and then decreasing again to 12 kDa at the tail; this variation was accompanied by the corresponding changes in the radii of gyration (R_g). This trend suggests a concentration-dependent post-elution effect, but it is unclear whether this is due to self-association or non-specific aggregation. Averaged peak-centre and peak-tail datasets for each protein demonstrate this shift to larger particles across the elution profile (Appendix Figure 4.2), and the corresponding calculated properties are reported in Appendix Table 4.2.

These results differ from the observations made previously by Maekawa *et al* (12), who, in addition to the crystal structure, presented *in vitro* data from analytical SEC which suggested the MLA10$^{5\text{-}120}$ protein behaved as a dimer in solution. To ensure the two sets of recombinant samples were consistent, we repeated these experiments on our own purified MLA10$^{5\text{-}120}$ protein. We found that our MLA10$^{5\text{-}120}$ does indeed elute at a volume similar to the 25 kDa chymotrypsin marker during SEC (Appendix Figure 4.3). This was also the case for Sr33 and Rx, and the discrepancy with the MALS data suggests that the elution time is influenced by protein properties other than size (Appendix Figure 4.3).

Our data shows that Sr33$^{6\text{-}120}$ does not exhibit detectable self-association *in vitro* at the concentrations tested. Some limited self-association was observed for MLA10$^{5\text{-}120}$ and Rx$^{1\text{-}122}$ during

SEC-SAXS at high protein concentration, but this was not apparent by SEC-MALS. The data agrees that the predominant species for both proteins is monomeric.

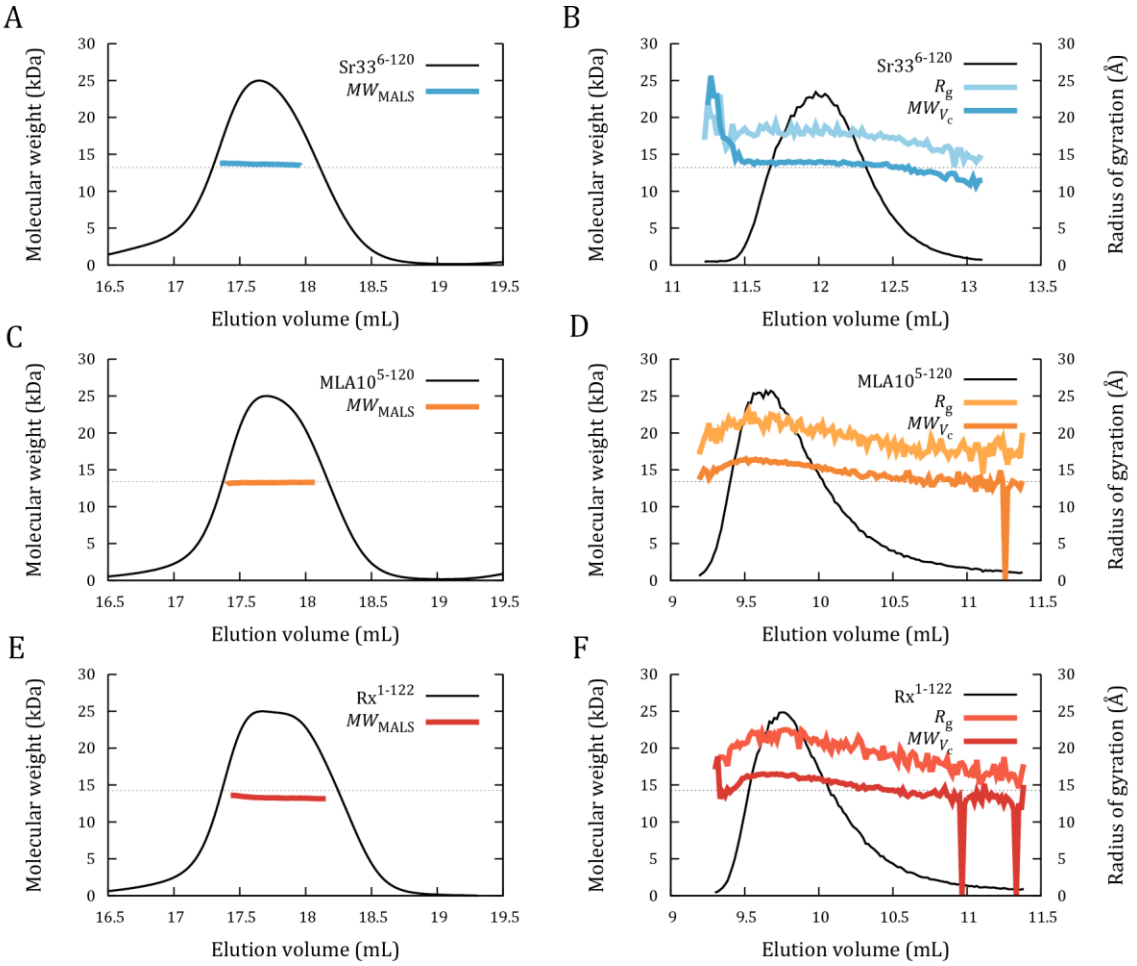


Figure 4.2 Sr33$^{6\text{-}120}$, MLA10$^{5\text{-}120}$ and Rx$^{1\text{-}122}$ are largely monomeric in solution. (**A–C**) Molecular weight calculations based on SEC-MALS analysis for Sr33$^{6\text{-}120}$ (**A, B**), MLA10$^{5\text{-}120}$ (**C, D**) and Rx$^{1\text{-}122}$ (**E, F**)**.** For all proteins, the solid black lines represent the normalized refractive index trace (arbitrary units, y-axis) for proteins eluted from an in-line Superdex 200 10/300 column [GE]. Coloured lines under the peaks correspond to the averaged molecular weight (MW; right-hand y-axis) distributions across the peak as determined by MALS. (**A, C, E**) Evolution of particle R_g and molecular weight during in-line SEC-SAXS with a WTC-030S5 column, for Sr33$^{5\text{-}120}$ (**B**), MLA10$^{5\text{-}120}$ (**D**) and Rx$^{1\text{-}122}$ (**F**). For all proteins, the trace of zero-angle intensity, $I(0)$, is plotted as a black line, whereas both R_g and MW_{VC} are plotted as coloured lines against the right-hand *y*-axis. R_g is shown with experimental errors at 1σ in lighter colours. The predicted molecular weight of each construct is shown as black dotted lines. The use of a 2 mL pre-column for Sr33$^{6\text{-}120}$ shifts that peak by the corresponding volume.

4.6.3 *Crystal structure re-analysis of MLA10$^{5-120}$*

The crystal structure of MLA10$^{5-120}$ reported by Maekawa *et al* (12) shows a helix-loop-helix structure that forms a dimer through a large interface. There is an unusual chain-break in helix α2, where five residues (91–95) could not be modelled due to the absence of electron density (12). This size of the interface and favourable interactions between residues from different monomers led the authors to suggest that the association would likely be permanent. By contrast, our solution studies of the equivalent protein do not support the formation of a stable dimer. To investigate possible reasons for the crystallization of the dimer, we attempted to crystallize the analogous protein construct under different experimental conditions to those previously defined. We found that MLA10$^{5-120}$ formed crystals in a large number of chemically diverse conditions. As the published MLA10$^{5-120}$ structure was crystallized at low pH (4.6) and high salt (2.0 M sodium formate), we decided to investigate crystal-forming conditions that involved a different type of precipitant and more neutral pH (MIB buffer pH 7.0 (42), 25% PEG 1500). These crystals diffracted X-rays to ~2.0 Å. However, attempts to determine the structure by molecular replacement, using the structures of the CC domains of either MLA10 (49) or Rx (21) as search models, were not successful. The protein was subsequently expressed to incorporate seleno-methionine instead of methionine, and the structure was determined by single-wavelength anomalous diffraction (SAD) (for further details see Methods and Appendix Table 4.3).

We were able to model all the protein residues (5–120), including the previously undefined region, residues 91–95, (Figure 4.3, Appendix Table 4.3), and the resolved MLA10$^{5-120}$ structure resembles closely the structure described previously (49). The dimeric interface identical to the one described previously (49) was also found in our crystal structure. Therefore, although we were able to obtain independently a very similar crystal structure to that of Maekawa and colleagues (49), the crystal structure remains inconsistent with our observations of the molecular weight calculations of MLA10$^{5-120}$ in solution.

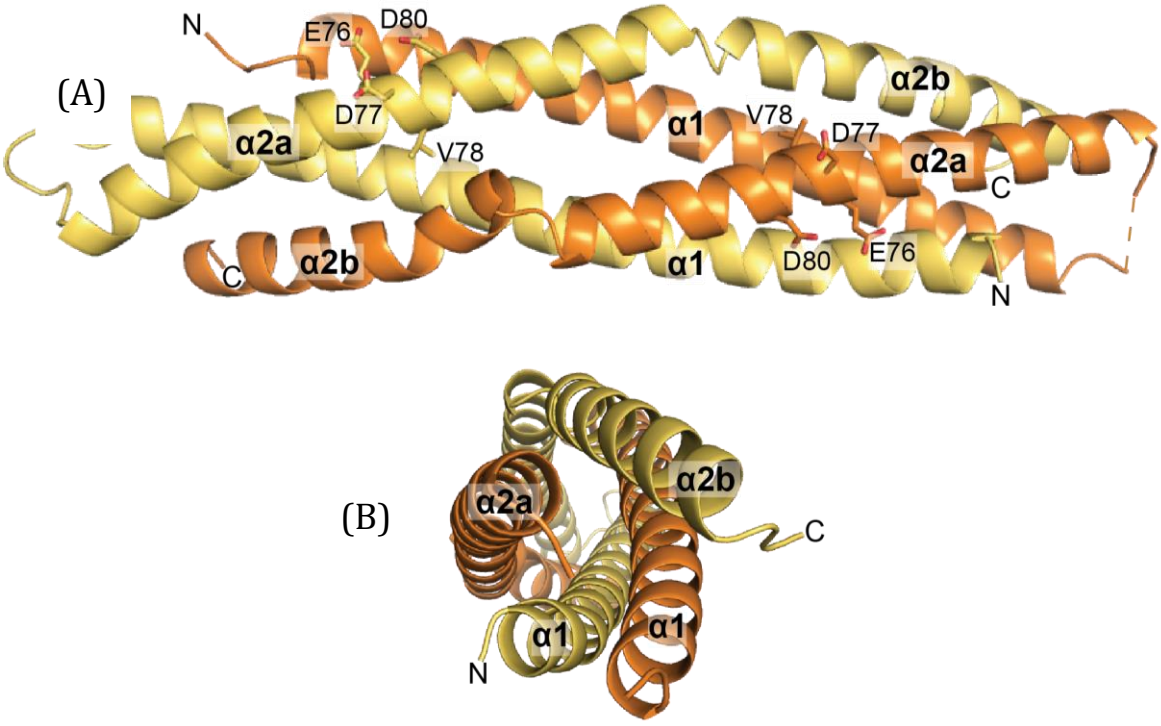


Figure 4.3 Cartoon representation of the crystal structure of MLA10$^{5\text{-}120}$. (**A**) Crystal structure of two MLA10$^{5\text{-}120}$ monomers (yellow and orange) in cartoon representation, with the individual helices and N- and C-termini labelled. Five residues (91–95) could not be modelled due to the absence of electron density, and were drawn by dashed line in orange between helix α1 and α2. The conserved EDVID motif is shown in stick representation. (**B**) 90° rotation around the y-axis of (A). Structural statistics found in Appendix Table 4.3.

4.6.4 *CC domains form compact and globular structures in solution*

SEC-SAXS data also contains information on the particle shape in solution, and we proceeded to investigate the solution structures of the CC-domain proteins. The scattering from the tail fractions of all three proteins is indistinguishable within the experimental errors (Figure 4.4A), and their calculated properties (Appendix Table 4.2) and real-space distributions also agree (Figure 4.4B). Moreover, the experimental data is consistent with the predicted scattering of the four-helix bundle arrangement observed in the structures of Sr33$^{6\text{-}120}$ and Rx$^{1\text{-}122}$, and is inconsistent with the structures of both the dimer and individual protomers from the MLA10$^{5\text{-}120}$ crystal structure (Figures 4.4C–E, Appendix Table 4.4). These fits can also be visualized if *ab initio* reconstructions from the SAXS data are superimposed onto the corresponding high-resolution structures (Figure 4.4F). All

three data-sets yield compact, globular shape envelopes, into which the NMR structure of $Sr33^{6-120}$ and the crystal structure of Rx^{1-122} can be docked with good agreement. These envelopes are clearly smaller than the extended conformation seen in the $MLA10^{5-120}$ crystals. The datasets collected at the peak centres are affected by self-association, rendering these unsuitable for modelling, and they do not compare well to any individual CC structure (Appendix Table 4.4). Collectively, the data suggests that the monomeric forms of $Sr33^{6-120}$, $MLA10^{5-120}$ and Rx^{1-122} all adopt compact, globular conformations, which are consistent with the four-helix bundle structure.

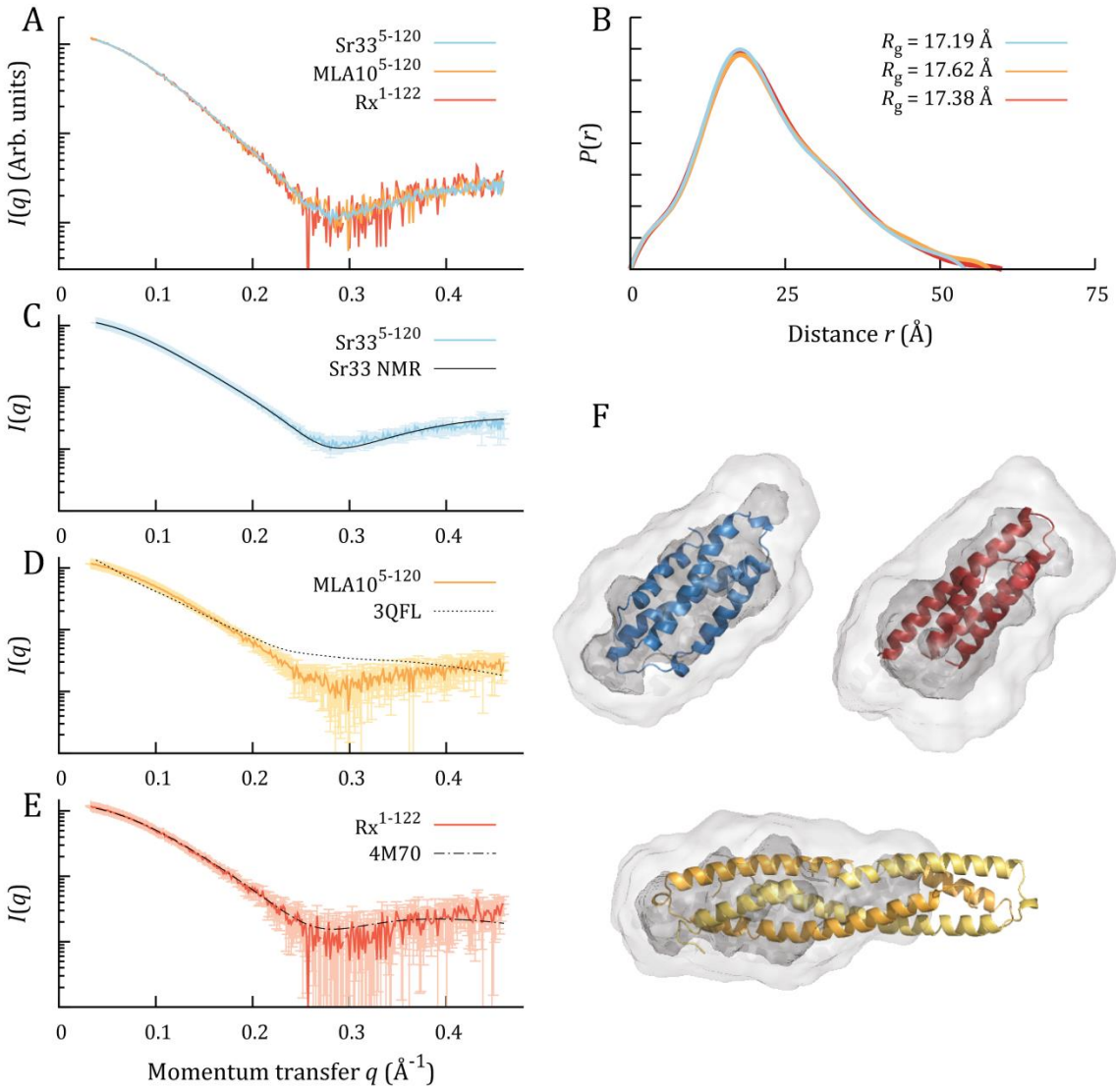
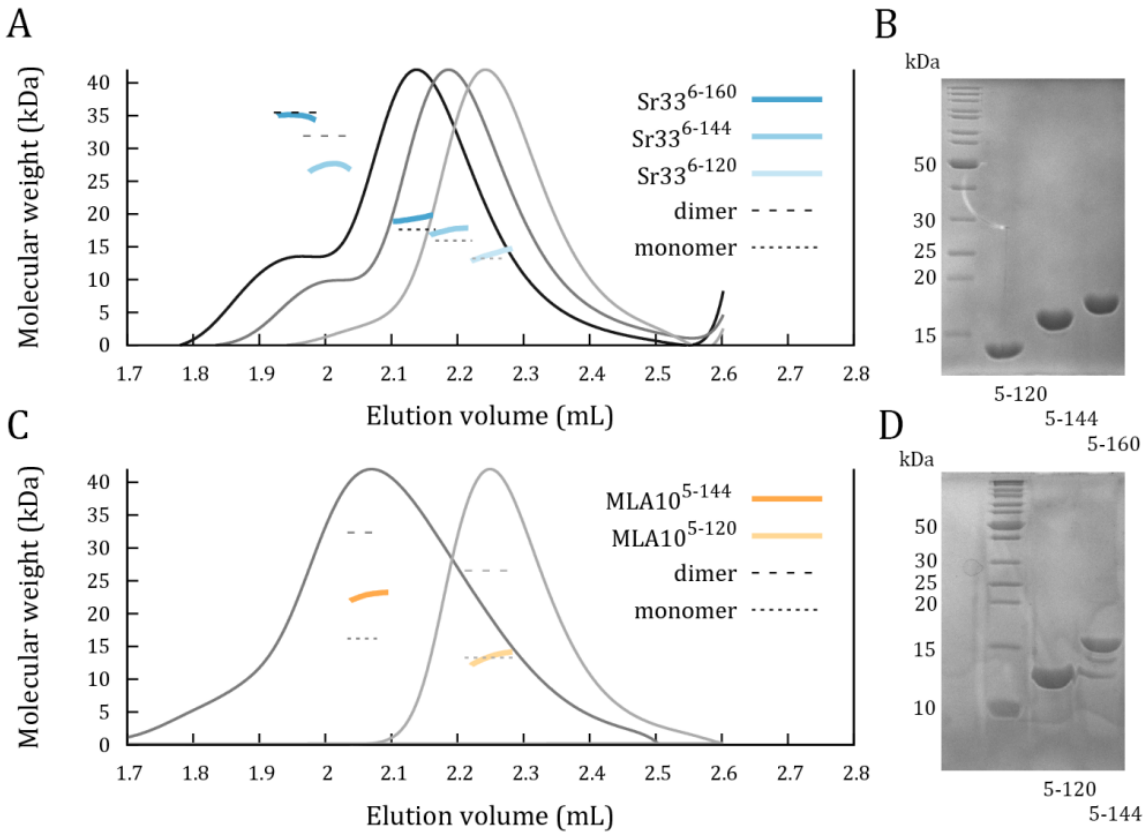


Figure 4.4 SAXS data from monomeric fractions of Sr33$^{6\text{-}120}$, MLA10$^{5\text{-}120}$ and Rx$^{1\text{-}122}$ is consistent with compact, globular particles. (**A**) Data-sets from SEC-SAXS are shown as coloured lines, with the MLA10$^{5\text{-}120}$ and Rx$^{1\text{-}122}$ data scaled to overlay with the Sr33$^{6\text{-}120}$ data. (**B**) Normalized distance distribution functions, *P*(*r*), are shown as coloured lines matching the scattering curve from which they were calculated. All distributions have been scaled to the maxima of the highest peaks. (**C**–**E**) SEC-SAXS data-sets again plotted as coloured lines. Experimental errors are displayed at 1σ in lighter colours. The theoretical scattering predicted from each 3D structural models is shown as a black line against the corresponding data-set. (**F**) The first member of the Sr33$^{6\text{-}120}$ NMR ensemble (blue), the Rx$^{1\text{-}122}$ crystal structure (red) and the dimeric MLA10$^{5\text{-}120}$ crystal structure (yellow) are shown in cartoon representation, docked into *ab initio* envelopes calculated from their respective scattering data-sets. *Ab initio* models are shown in transparent surface representation, with the average model from 16 independent runs shown in light grey and the filtered model in darker grey. Divergence between the MLA10 crystal structure and the SAXS envelope is clear.

4.6.5 *Extended CC domain fragments of Sr33 and MLA10 CC domains show an increased propensity to self-associate*

According to secondary-structure predictions (50), residue 120 of both Sr33 and MLA10 resides within a helix that is predicted to extend until residue 138 (Appendix Figure 4.4). Recent *in planta* results indicate that the residues between 120–160 are required for signalling and self-association (51). Consequently, we expressed in *E. coli*, purified and prepared these larger fragments of the CC domains for analysis by SEC-MALS. A modified purification buffer system was necessary for purification (see Methods). We were able to purify $Sr33^{6-144}$, $Sr33^{6-160}$ and $MLA10^{5-144}$ to homogeneity (Figure 4.5B, D), but the $MLA10^{5-160}$ fragment was largely insoluble.

SEC-MALS revealed an increasing propensity for self-association in the longer CC domain fragments (Figures 4.5A, C). The behaviour of $Sr33^{6-120}$ and $MLA10^{5-120}$ remained consistent with that reported in the previous experimental setup (Figures 4.2A, C, E). For Sr33, an earlier peak with a molecular weight near the expected dimeric size was apparent for $Sr33^{6-144}$ and more pronounced for $Sr33^{6-160}$ (Figure 4.5B). A larger peak at the expected monomeric size was still present for both. $MLA10^{5-144}$ showed a single peak that was much more extended and asymmetric than that observed in $MLA10^{5-120}$, and with a molecular weight 40% higher than the expected monomeric weight. These experiments demonstrate that extended CC constructs of both Sr33 and MLA10 undergo self-association *in vitro*, resulting in a mix of monomeric and dimeric species. Despite this, the monomeric forms remain the predominant forms in all constructs.



E

Construct	Theoretical MW (kDa)		Experimental MW (kDa)	
	Monomer	Dimer	Shoulder peak	Main peak
$Sr33^{6-120}$	13.12	26.24	-	13.9
$Sr33^{6-144}$	15.94	31.88	27.3	17.5
$Sr33^{6-160}$	17.65	35.30	35.0	19.3
$MLA10^{5-120}$	13.28	26.56	-	13.4
$MLA10^{5-144}$	16.17	32.34	-	22.8

Figure 4.5 Solution studies of CC domains with extended sequences of Sr33 and MLA10. (**A, C**) Molecular-weight calculations from SEC-MALS analysis for $Sr33^{6-120}$, $Sr33^{6-144}$ and $Sr33^{6-160}$ (**A**); and $MLA10^{5-120}$ and $MLA10^{5-144}$ (**C**)**.** Solid-grey, dark-grey and black lines represent the refractive index for the three proteins, respectively, when eluted from an in-line Superdex 200 5/150 GL column [GE]; these have been normalized to the height of the major peak for clarity. Dotted lines indicate the predicted molecular weights of both monomeric and dimeric species, and coloured lines show the experimental molecular-weight distributions as determined by MALS. (**B, D**) Coomasie blue-stained SDS-PAGE of purified (left to right) $Sr33^{6-120}$, $Sr33^{6-144}$ and $Sr33^{6-160}$ (**B**); and $MLA10^{5-120}$ and $MLA10^{5-144}$ (**D**) proteins, used in solution studies. (**E**) Summary of predicted and experimental average molecular weights for all constructs as determined by MALS.

4.6.6 *Defining the minimal CC-domain signalling unit in Sr33, MLA10 and Sr50*

In light of our findings and those by Cesari and colleagues (49), we set out to identify the minimal N-terminal fragment necessary for the signalling function of MLA10, Sr33 and Sr50, a close homolog of Sr33 isolated from Rye (52). We generated six truncations of MLA10, Sr33 and Sr50 N-terminal domains within the region equivalent to MLA10 residues 120–160, at positions surrounding the predicted end of the α4 helix (1–130, 1–135, 1–141, 1–142, 1–144 and 1–148) (Figure 4.5A). These N-terminal fragments were transiently expressed in *N. benthamiana* under the control of the 35S promoter and fused to a C-terminal HA tag. These results showed that fragments truncated at, or beyond, the equivalent of MLA10 residue 142 were auto-active and induced cell death while shorter fragments were inactive (Figure 4.6A). The MLA10$^{1\text{-}160}$:CFP, Sr33$^{1\text{-}160}$:CFP and Sr50$^{1\text{-}163}$:CFP constructs were used as positive controls and triggered cell death (49). Immunoblotting showed that all fusion proteins were properly expressed (Appendix Figure 4.5A). Taken together, these results identify the minimal N-terminal cell death signalling domains of MLA10, Sr33 and Sr50 as extending to amino-acid position equivalent to 142 in MLA10.

4.6.7 *Auto-active fragments of MLA10, Sr33 and Sr50 self-associate in planta*

To investigate whether association *in planta*, indicated by induction of cell death, was correlated with a measure of *in vitro* association, we performed co-immunoprecipitation (co-IP) experiments using fragments of MLA10, Sr33 and Sr50 equivalent to MLA10 1–141, 1–142 and 1–144 fused to CFP or HA tags (Figure 4.6B). CFP-fused CC fragments were expressed (Appendix Figure 4.5B) and displayed equivalent cell death activity as corresponding HA-tagged fragments (Appendix Figure 4.5C). Proper expression of all proteins in the input was verified by immunoblotting using anti-GFP and anti-HA antibodies (Figure 4.6B). CFP-fused proteins were enriched after immunoprecipitation with anti-GFP beads and HA-fused auto-active fragments of MLA10, Sr33 and Sr50 co-precipitated with their respective CFP-fused fragment. However, they did not co-precipitate, or to a weaker background extent, with the RGA4$^{1\text{-}171}$:CFP construct (53) that was used as a negative control for binding specificity. Interestingly, self-association was highly reduced in the case of inactive CC domains. The Sr50$^{1\text{-}163}$ construct was used as a positive control for *in planta* self-association (49). Taken together, these results show that auto-active CC fragments of MLA10, Sr33, and Sr50 form specific homo-complexes in planta while inactive fragments do not self-associate, or do so to a limited extent. Therefore, these data indicate a clear correlation between self-association and *in planta* signalling activity.

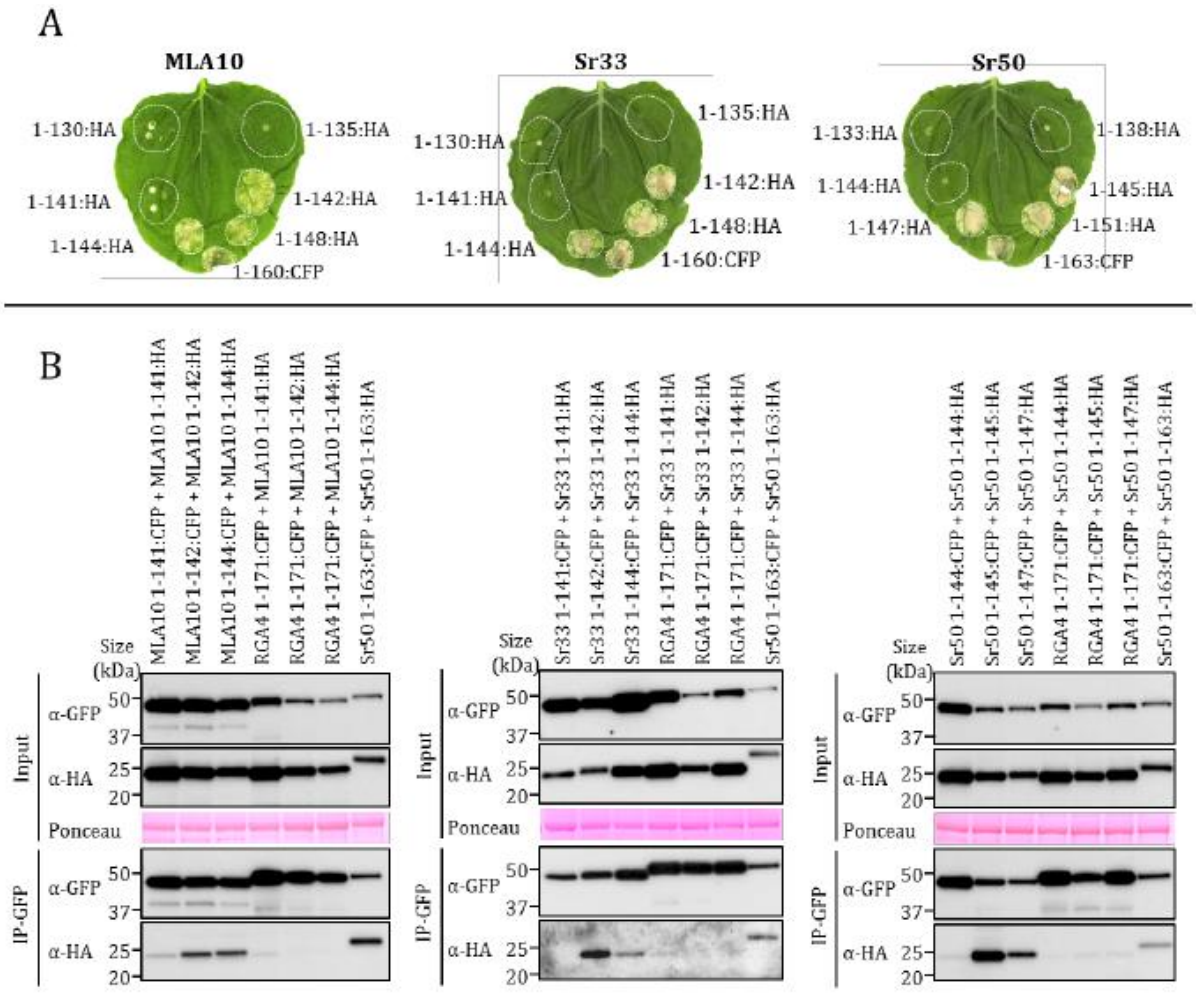


Figure 4.6 Minimal autoactive domains of MLA10, Sr33, and Sr50 self-association *in planta*. (**A**) The indicated fragments of the MLA10, Sr33, and Sr50 proteins fused to HA or CFP were transiently expressed in *N. benthamiana* leaves by *A. tumefaciens* infiltration. The autoactive MLA10$^{1\text{-}160}$:CFP, Sr33$^{1\text{-}160}$:CFP, and Sr33$^{1\text{-}163}$:CFP constructs were used as positive controls. Cell death was visualised five days after infiltration. Equivalent results were obtained in 3 independent experiments. (**B**) The indicated fragments of MLA10, Sr33, and Sr50 fused to CFP or HA tags were transiently expressed in *N. benthamiana* leaves. Proteins extracted 20 hours after infiltration were analysed by immunoblotting with anti-HA (α-HA) and anti-GFP antibodies (α-GFP). Proteins were immunoprecipitated with anti-GFP beads (IP-GFP) and analysed by immunoblotting with anti-GFP and anti-HA antibodies. The CC domain of RGA4 fused to the CFP was used as a control for specificity, and the Sr50$^{1\text{-}163}$ construct was used as a positive control in each experiment (45). Ponceau staining of the RuBisCO large subunit shows equal protein loading in the input.

4.7 Discussion

4.7.1 *Conservation of a plant NLR CC domain structure*

Prior to this study, the partial structures of two plant CC domains had been reported. The four-helix bundle structure of the CC domain from the potato NLR protein Rx is 'strikingly different' (21) from the dimeric helix-turn-helix structure of the CC domain from the barley NLR protein MLA10 (12). Given the low sequence identity between MLA10 and Rx within the structured region (~18%) and the fact that Rx^{1-122} was crystallized in complex with a fragment of its co-factor, it was perhaps not surprising that the two proteins would adopt different conformations. Hao and co-workers (21) noted that the identification of any representative structure for the class would require characterization of further CC structures. Interestingly, such variation has not been observed among TIR-domains, the other class of plant NLR N-terminal domains. These have so far been found to adopt a conserved fold even between members with low sequence identities, raising the possibility that the plant NLR CC domains may adopt more diverse roles, functions and mechanisms of signalling than the TIR domains.

However, it is surprising that the NMR structure of the $Sr33^{6-120}$ CC domain presented here revealed a compact four-helix bundle similar to the crystal structure of the Rx CC domain (18% sequence identity) (21), rather than the dimeric arrangement in the crystal structure of MLA10 (~82% sequence identity) (49). Biophysical characterization using SEC-MALS and SAXS confirmed predominantly monomeric behaviour in solution for $Sr33^{6-120}$, Rx^{1-122}, and unexpectedly, also $MLA10^{5-120}$. Some limited self-association was observed in SEC-SAXS experiments for Rx^{1-122} and $MLA10^{5-120}$, but not what would be expected for an obligate dimer suggested by the crystal structure for $MLA10^{5-120}$ (12). In addition, the monomeric fractions of all three samples were indistinguishable by SEC-SAXS shape analysis, and were consistent with the 4-helix bundle structures of the $Sr33^{6-120}$ and Rx^{1-122}, but not the dimeric $MLA10^{5-120}$ crystal structure. These data suggest that plant NLR CC domains may have conserved structures, as is the case for TIR domains.

4.7.2 *Self-association and CC domain signalling*

Highly relevant to this work is the recent discovery that constructs of Sr33 and MLA10 comprising residues 1–120 of the CC domain do not co-IP *in planta*, consistent with our finding that $Sr33^{6-120}$ and $MLA10^{5-120}$ do not self-associate in solution. More importantly, these truncated constructs also did

not induce a cell-death phenotype when transiently expressed in tobacco. Longer constructs comprising residues 1–160 were capable of both *in planta* co-IP and cell-death activity.

Our solution studies using these longer active constructs showed that the inclusion of additional residues at the C-terminus (MLA10$^{5\text{-}144}$, Sr33$^{6\text{-}144}$ and Sr33$^{6\text{-}160}$) facilitates self-association, corroborating the link between self-association and signalling for these proteins. It is, however, important to note that even the longest constructs retained a large fraction of monomeric protein in solution, suggesting that both monomer and dimer forms are present.

Building upon this, we demonstrated using transient assays in tobacco that for Sr33, MLA10 and the related wheat NLR Sr50, the minimal functional unit for cell-death signalling extends to a position that coincides with the predicted end of the last α-helix within the CC domain. In addition, co-IP experiments demonstrated that any constructs that were capable of causing cell death could also self-associate when expressed *in planta*. Collectively, these results establish a correlation between self-association and biological activity and suggest that self-association is the switch regulating cell-death induction.

4.7.3 *The coiled-coil dimer*

It is important to address the tendency of MLA10$^{5\text{-}120}$ to crystallize in a dimeric conformation, as well as the conflicting evidence for such a dimer in solution. Maekawa and colleagues suggested MLA10$^{5\text{-}120}$ existed as a dimer in solution, based on the crystal structure, the slow migration by SEC, the appearance of cross-linked bands following incubation with a chemical crosslinking agent, and the insolubility of recombinant protein after the putative dimer interface was disrupted by mutation (12). We were able to independently confirm the results of crystallography, SEC and cross-linking, suggesting the conflicting observations require explanation. Migration by SEC is often used to estimate molecular weights of proteins by comparison to protein standards of known molecular weights; however, the migration is influenced by particle shape, flexibility and composition. By contrast, determination of molecular weight using MALS and SAXS techniques relies on absolute measurements, and while there is no ideal method, these approaches provide a more robust and quantitative means of molecular-weight measurement than migration by SEC (35,54,55). In the case of MLA10$^{5\text{-}120}$, these methods show only a small degree of transient self-association. Consistent with these data, MLA10$^{5\text{-}120}$ remains predominantly monomeric even after extended incubation with a crosslinking agent and the observed behaviour is likely more indicative of weak self-association than an obligate interaction.

Similarly, the process of crystallization selects for states that will promote a highly ordered arrangement in the crystal. The comparison pf the dimer observed in the crystals to the four-helix bundle structures suggests that the monomers may have undergone a domain swap in the crystallization solution. Figure 4.1 shows that it is possible to superimpose two copies of the $Sr33^{6-120}$ four-helix bundle side-by-side onto the MLA10 dimer. Repositioning α1 and α4 in $Sr33^{6-120}$ to form a continuous helix would regenerate the helix-loop-helix observed in the $MLA10^{5-120}$ dimer, while retaining all the same internal hydrophobic contacts. Importantly, this also explains the insolubility of the interface-disrupting mutants. The residues that form the dimeric interface also form the hydrophobic core of the four-helix bundle monomer, and the same mutations would also disrupt the monomeric fold and render the protein insoluble. Domain swapping is quite common (56), and may be caused by the dynamic behaviour of the protein that could reflect conformational transitions that have functional significance in signalling.

Our results demonstrate that extensions to residue 144 in both Sr33 and MLA10 are necessary for consistently measurable self-association in solution. Importantly, this self-association is highly correlated with signalling activity. This tight correlation suggests that it is the dimeric form that is responsible for cell death signalling of the CC domain. However, the nature of this dimer is not clear. While it remains plausible that the $MLA10^{5-120}$ crystal structure captures some part of the activated signalling dimer, it is also possible that the additional C-terminal residues promote self-association of the protein in the four-helix-bundle conformation instead. Based on our data, we favour the latter of these explanations. Importantly, the measured molecular weight of $MLA10^{5-144}$ from SEC-MALS analysis is between the expected sizes of the monomeric and dimeric forms of the protein. Interestingly, $Sr33^{6-160}$ shows evidence of such a separated dimeric peak, albeit only a small proportion of the protein is dimeric (Appendix Figure 4.3B). Clearly, a full structural exploration of the longer, active constructs will be necessary to reveal the conformation of these dimers and the protein-protein interfaces that control their formation.

4.7.4 *Mechanism of signalling by NLR CC domains*

It has been hypothesized that the self-association of the TIR domains post-activation is positively regulated via self-association of other domains from the full-length plant NLR (57). There is a precedent for self-association of NB and LRR domains in animal NLRs (17-19), although this has yet to be observed structurally in plant NLRs. We propose a similar model of signalling for the CC-NLRs, in which the transient self-association of the CC domain is stabilized by the full-length NLR in order

to achieve the activated state. These associations would presumably facilitate the recruitment of downstream signalling molecules, as is the case in animal NLRs (17-19,58) and Toll-like receptors (59).

Our study demonstrates that both distantly and closely related CC-NLR proteins have structurally similar CC domains, reconciling previously conflicting data and models of activation of this important domain. We show that self-association of CC domains is highly correlated with cell-death activity in the MLA10 and Sr33 CC-NLR proteins, and we define residues comprising a minimal-functional unit both biophysically and *in planta*. Collectively, this data redefine our structural understanding of the CC domains from CC-NLR proteins. It is anticipated that this work will provide a platform for further structural studies of the full-length CC domains and illuminate the molecular details of the homo-interactions that are associated with defence signalling.

4.8 Acknowledgments

This research was supported by the Australian Research Council (ARC) Discovery Projects DP120100685 and DP160102244. B.K. is an NHMRC Research Fellow (1003325 and 1110971). M.M. is an ARC Future Fellow (FT110100925). S.J.W. is funded by ARC DECRA (DE160100893). We acknowledge the use of the University of Queensland Remote Operation Crystallization and X-ray Diffraction Facility (UQ-ROCX). The SAXS and X-ray diffraction data collection was undertaken on the SAXS-WAXS and MX beamlines at the Australian Synchrotron, respectively. We thank the beamline scientists at the Australian Synchrotron for their advice and support. We thank Sambasivam Periyannan and Evans Lagudah for access to the *Sr33* cDNA. We declare that we have no conflict of interest.

4.9 Chapter 4 appendix

Appendix Table 4.1 NMR structure statistics[a]

Experimental restraints[b]		
	Inter-proton distance restraints	
	Intra-residue	580
	Sequential	186
	Medium-range (i–j < 5)	293
	Long-range (i–j > 5)	249
	Dihedral-angle restraints	214
	Total number of restraints per residue	12.96
RMSD from mean coordinate structure (Å)[c]		
	Backbone atoms (residues 6–89 & 98–110)	0.93 ± 0.21
	All heavy atoms (residues 6–89 & 98–110)	1.31 ± 0.19
Stereochemical quality[d]		
	Residues in most favoured Ramachandran region (%)	93.1
	Ramachandran outliers (%)	0 ± 0

[a]All statistics are given as mean ± SD.

[b]Only structurally relevant restraints, as defined by CYANA, are included.

[c]Mean r.m.s. deviation calculated over the entire ensemble of 20 structures.

[d]As reported by CYANA (31,32).

Appendix Table 4.2 Properties derived from averaged SAXS datasets

Protein	Fraction	Elution range (mL)	$I(0)_{Guin}$ (cm^{-1})	$I(0)_{P(r)}$ (cm^{-1})	$R_{g\ Guin}$ (Å)	$R_{g\ P(r)}$ (Å)	MW_{Vc} (kDa)	MW_{Porod} (kDa)
$Sr33^{5-120}$	Centre	11.77 – 12.31	4.00 e^{-2}	4.09 e^{-2}	17.73	18.81*	14.1	16.3
	Tail	12.35 – 12.69	1.28 e^{-2}	1.28 e^{-2}	17.08	17.19	13.5	13.9
$MLA10^{5-120}$	Centre	9.29 – 10.04	3.69 e^{-2}	3.76 e^{-2}	20.81	22.27*	16.1	25.0 †
	Tail	10.42 – 11.29	0.46 e^{-2}	0.46 e^{-2}	17.72	17.62	13.7	15.5
Rx^{1-122}	Centre	9.48 – 10.07	3.95 e^{-2}	4.01 e^{-2}	21.35	22.63*	16.3	25.0 †
	Tail	10.65 – 11.32	0.31 e^{-2}	0.31 e^{-2}	17.10	17.38	13.7	15.1

* Values for $R_{g\ P(r)}$ that differ from $R_{g\ Guin}$ by greater than 5%

† Values for MW_{Porod} that differ from MW_{Vc} by greater than 15%

The theoretical monomeric molecular weights of $Sr33^{5-120}$, $MLA10^{5-120}$ and Rx^{1-122} are 13.2 kDa, 13.4 kDa and 14.3 kDa, respectively.

Appendix Table 4.3 Crystallographic table for MLA10[5-120]

Data processing		
Space group	P 2 21 21	P 1
a, b, c (Å)	30.87 87.56 92.56	
α, β, γ (°)	90 90 90	
Resolution (Å)	46.28-2.1 (2.16–2.10)	
Rmeas(%) [a]	0.11 (1.94)	
Rpim(%) b	0.03 (0.51)	
$\langle I/\sigma(I)\rangle$	15 (1.8)	
$CC_{1/2}$ [c]	0.99 (0.89)	
Completeness (%)	100 (100)	
Multiplicity	14.1 (14.4)	
Wilson plot B (Å^2)	44.7	
Observations	216711 (18084)	
Unique reflections	15392 (1253)	
Anomalous completeness	100 (100)	
Anomalous multiplicity	7.7 (7.7)	
DelAnom correlation between	0.471 (-0.027)	
Mid-Slope of Anom Normal	1.087	
Estimate of maximum resolution for significant anomalous signal = 3.59A, from CCanom > 0.15		
Refinement		
Rwork (%)		24.5
Rfree (%)		27.8
Average B-factor (Å^2)		56.6
R.m.s deviations		
Bond lengths (Å)		0.0101
Bond angles (°)		1.08
Ramachandran plot (%) [d]		
Favoured		97.95
Allowed		1.71
Outliers		0.34

[a] $R_{meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith measurement of an equivalent reflection with indices hkl.

[b] $R_{pim} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

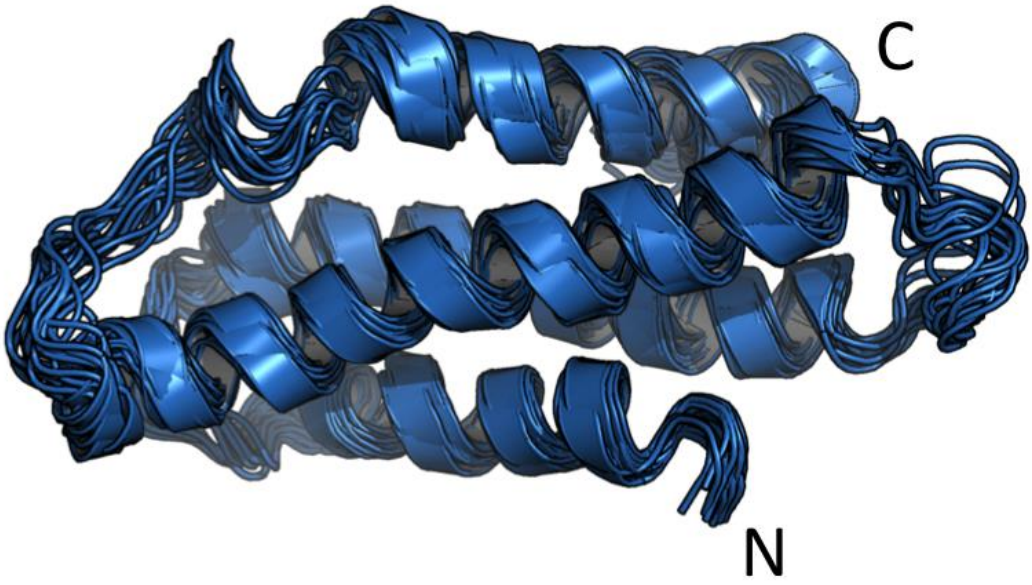
[c] Calculated with the program Aimless (45)

[d] As calculated by MolProbity (60)

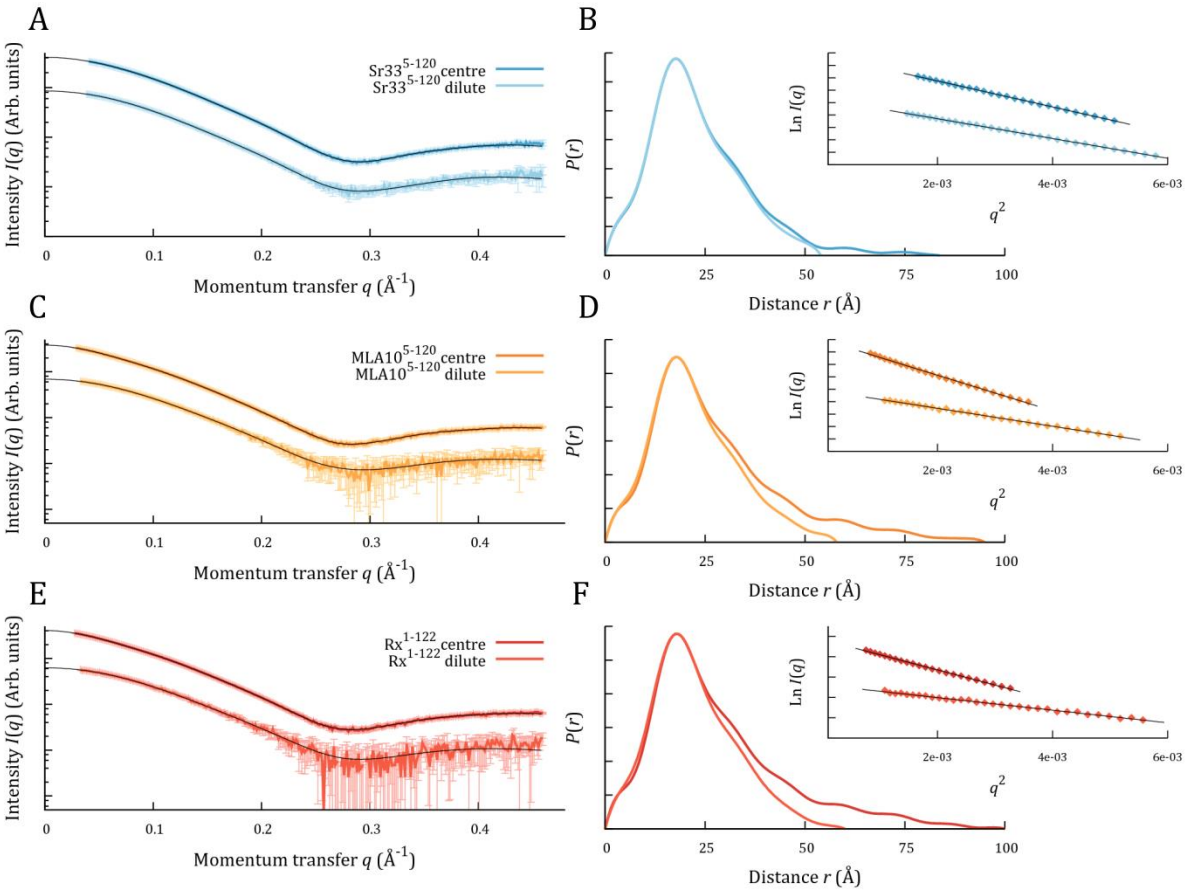
NB: Values within parentheses indicate the highest resolution bin.

Appendix Table 4.4 Goodness-of-fit (χ) scores for averaged SAXS datasets compared to structures

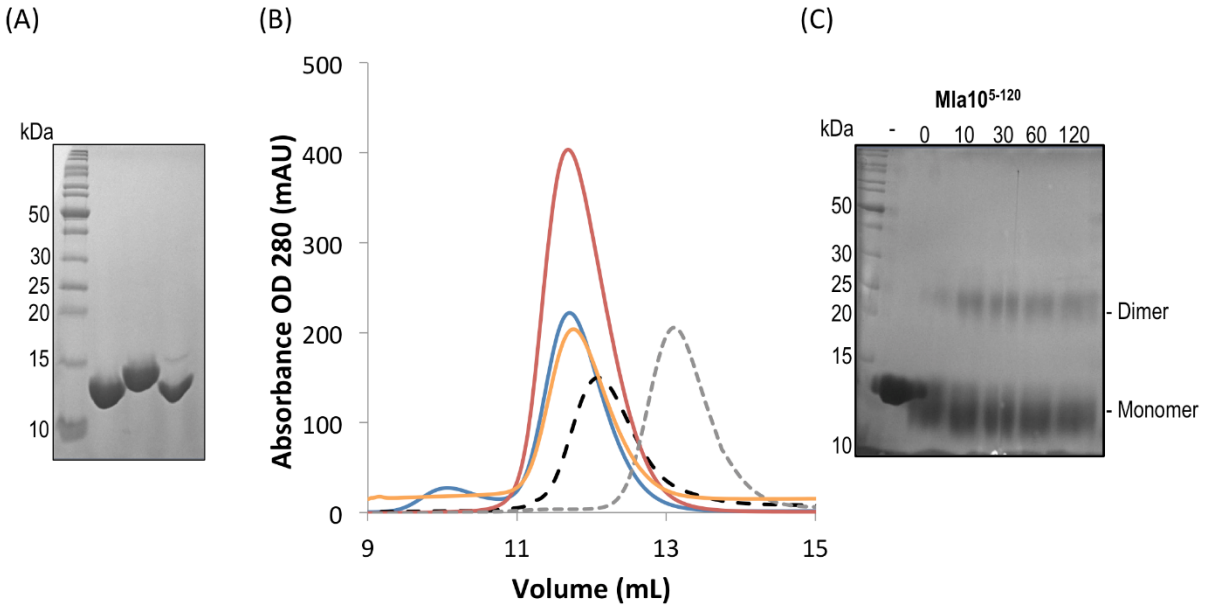
Sample	Fraction	Atomic structure (PDB ID)			
		$Sr33^{6-120}$	Rx^{1-122}	$MLA10^{5-120}$	
		NMR*	4M70	3QFL monomer	3QFL dimer
$Sr33^{5-120}$	Tail	0.67	1.63	12.03	8.41
$MLA10^{5-120}$	Tail	0.51	0.87	3.99	2.88
Rx^{1-122}	Tail	0.48	0.73	3.44	2.71
$Sr33^{5-120}$	Centre	2.47	7.36	31.62	20.54
$MLA10^{5-120}$	Centre	7.24	14.30	19.02	12.60
Rx^{1-122}	Centre	7.52	14.59	17.85	11.66



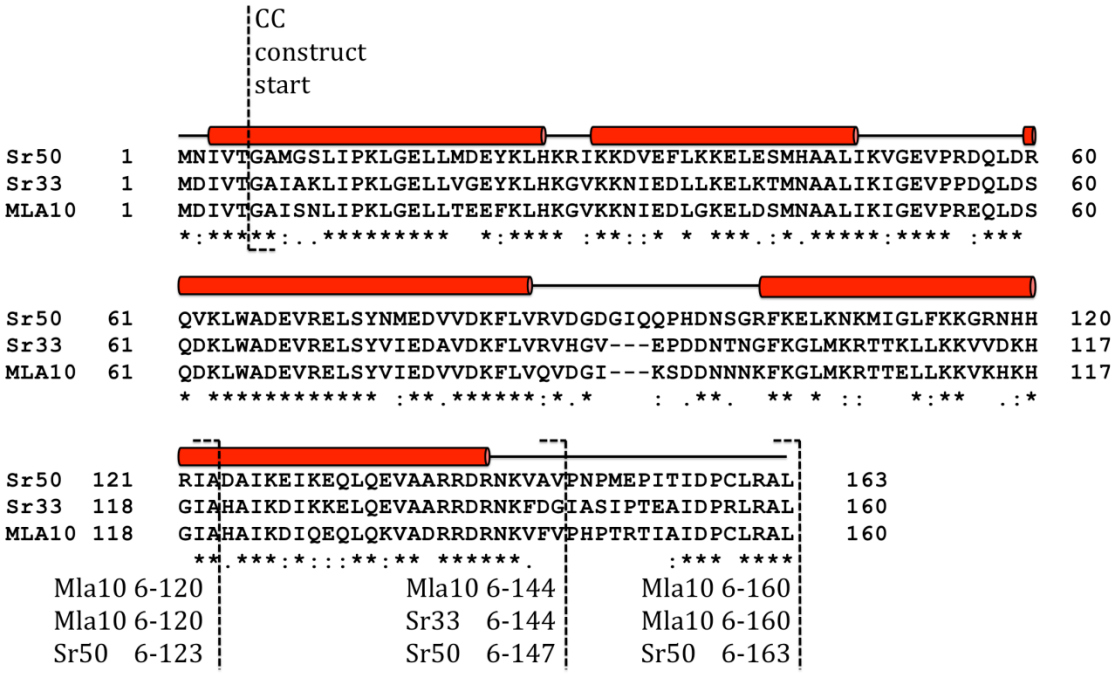
Appendix Figure 4.1 Ensemble of twenty super-imposed lowest-energy NMR structures of $Sr33^{6-120}$.



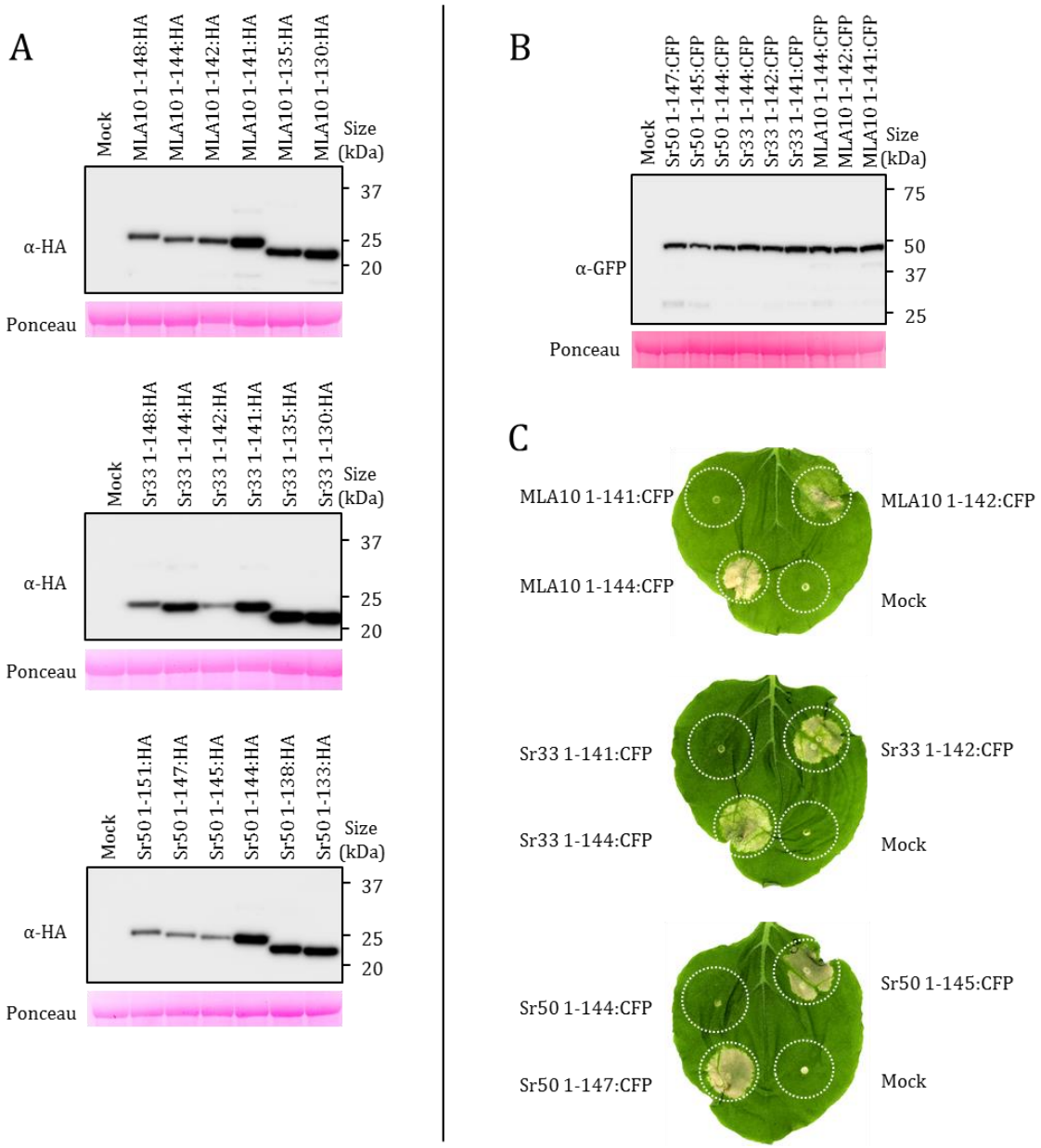
Appendix Figure 4.2 Analysis of scattering curves averaged over the peak centre and dilute fractions from SEC-SAXS. (**A**, **C**, **E**) Experimental datasets are plotted as coloured lines, with experimental errors displayed at 1σ in lighter colour. Solid black lines indicate the fit of the corresponding distance distribution. The datasets are arbitrarily offset along the *y*-axis for ease of visualization. (**B, D, F**) Normalized distance distribution functions, *P*(*r*), are shown as coloured lines matching the scattering curve from which they were calculated. *P*(*r*)s have been normalized to reciprocal-space zero-angle intensity. The Guinier regions of the datasets are shown in the insets, transformed as q^2 vs ln *I*(*q*). Individual data-points are plotted as coloured diamonds, and a linear regression fit to each is shown as a black line. The datasets are again offset in *y* for visualization.



Appendix Figure 4.3 Purification and SEC analysis (**A**) Coomassie blue-stained SDS-PAGE of purified proteins (left-right) Mla10$^{5\text{-}120}$, Rx$^{1\text{-}122}$ and Sr33$^{6\text{-}120}$ proteins. (**B**) Mla10$^{5\text{-}120}$ (orange), Rx$^{1\text{-}122}$ (red) and Sr33$^{6\text{-}120}$ (blue) were separated on a Superdex S75 10/300 size-exclusion chromatography column and compared with known standards chymotrypsin (25 kDa) and cytochrome c (12 kDa) (dashed lines). (**C**) Chemical crosslinking of Mla10$^{5\text{-}120}$. The protein was incubated with the cross-linker BS3 and sampled at time points 0, 10, 30, 60, 120 min. (-) represents the protein without BS3 added. The protein samples were separated by SDS-PAGE and Coomassie blue-stained.



Appendix Figure 4.4 Secondary structure prediction of Mla10, Sr33 and Sr50 CC domains from protein sequences using PSIPRED (61,62). Construct boundaries are marked with dashed lines.



Appendix Figure 4.5 HA and CFP-tagged CC fragments of MLA10, Sr33 and Sr50 are expressed in *N. benthamiana*. (**A**, **B**) Indicated proteins were extracted from transiently transformed *N. benthamiana* leaves 20 hours after infiltration and were analyzed by immunoblotting with anti-GFP or anti-HA antibodies. Ponceau staining of RuBisCO was used to verify equal protein loading. (**C**) Indicated constructs were transiently expressed in *N. benthamiana*. Cell death was visualized five days after infiltration.

4.10 Chapter 4 references

1. Jones, J. D. G., and Dangl, J. L. (2006) The plant immune system. *Nature Review* **444**, 323–329.
2. Dodds, P. N., and Rathjen, J. P. (2010) Plant immunity: towards an integrated view of plant-pathogen interactions. *Nature Review Genetics* **11**, 539–548.
3. Periyannan, S., Moore, J., Ayliffe, M., Bansal, U., Wang, X., Huang, L., Deal, K., Luo, M., Kong, X., Bariana, H., Mago, R., McIntosh, R., Dodds, P., Dvorak, J., and Lagudah, E. (2013) The gene Sr33, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. *Science* **341**, 786–788.
4. Tameling, W. I. L., Elzinga, S. D. J., Darmin, P. S., Vossen, J. H., Takken, F. L. W., Haring, M. A., and Cornelissen, B. J. C. (2002) The tomato R gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity. *Plant Cell* **14**, 2929–2939.
5. Williams, S. J., Sornaraj, P., DeCourcy-Ireland, E., Menz, R. I., Kobe, B., Ellis, J. G., Dodds, P. N., and Anderson, P. A. (2011) An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. *Molecular Plant-Microbe Interactions Journal* **24**, 897–906.
6. Bernoux, M., Burdett, H., Williams, S. J., Zhang, X., Chen, C., Newell, K., Lawrence, G., Kobe, B., Ellis, J. G., Anderson, P., and Dodds, P. N. (2016) Comparative analysis of the flax immune receptors L6 and L7 suggests an equilibrium-based switch activation model. *Plant Cell* **28**, 146–159.
7. Dodds, P. N., Lawrence, G. J., and Ellis, J. G. (2001) Six amino acid changes confined to the leucine-rich repeat β-strand/β-turn motif determine the difference between the P and P2 rust resistance specificities in flax. *Plant Cell* **13**, 163–178.
8. Ravensdale, M., Bernoux, M., Ve, T., Kobe, B., Thrall, P. H., Ellis, J. G., and Dodds, P. N. (2012) Intramolecular interaction influences binding of the flax L5 and L6 resistance proteins to their AvrL567 ligands. *PLoS Pathogens* **8**, e1003004–e1003004.
9. Hu, Z., Yan, C., Liu, P., Huang, Z., Ma, R., Zhang, C., Wang, R., Zhang, Y., Martinon, F., Miao, D., Deng, H., Wang, J., Chang, J., and Chai, J. (2013) Crystal structure of NLRC4 reveals its autoinhibition mechanism. *Science* **341**, 172–175.
10. Swiderski, M. R., Birker, D., and Jones, J. D. G. (2009) The TIR domain of TIR-NB-LRR resistance proteins is a signaling domain involved in cell death induction. *Molecular Plant-Microbe Interactions Journal* **22**, 157–165.
11. Bernoux, M., Ve, T., Williams, S., Warren, C., Hatters, D., Valkov, E., Zhang, X., Ellis, J. G., Kobe, B., and Dodds, P. N. (2011) Structural and functional analysis of a plant resistance

protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. *Cell Host Microbe* **9**, 200–211.

12. Maekawa, T., Cheng, W., Spiridon, L. N., Töller, A., Lukasik, E., Saijo, Y., Liu, P., Shen, Q.-H., Micluta, M. a., Somssich, I. E., Takken, F. L. W., Petrescu, A.-J., Chai, J., and Schulze-Lefert, P. (2011) Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host Microbe* **9**, 187–199.
13. Collier, S. M., Hamel, L.-P., and Moffett, P. (2011) Cell death mediated by the N-terminal domains of a unique and highly conserved class of NB-LRR protein. *Molecular Plant-Microbe Interactions Journal* **24**, 918-931.
14. Bai, S., Liu, J., Chang, C., Zhang, L., Maekawa, T., Wang, Q., Xiao, W., Liu, Y., Chai, J., Takken, F. L. W., Schulze-Lefert, P., and Shen, Q.-H. H. (2012) Structure-function analysis of barley NLR immune receptor MLA10 reveals its cell compartment specific activity in cell death and disease resistance. *PLoS Pathogens* **8**, 21–25.
15. Williams, S. J., Sohn, K. H., Wan, L., Bernoux, M., Sarris, P. F., Segonzac, C., Ve, T., Ma, Y., Saucet, S. B., Ericsson, D. J., Casey, L. W., Lonhienne, T., Winzor, D. J., Zhang, X., Coerdt, A., Parker, J. E., Dodds, P. N., Kobe, B., and Jones, J. D. G. (2014) Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science* **344**, 299–303.
16. Césari, S., Kanzaki, H., Fujiwara, T., Bernoux, M., Chalvon, V., Kawano, Y., Shimamoto, K., Dodds, P., Terauchi, R., and Kroj, T. (2014) The NB-LRR proteins RGA4 and RGA5 interact functionally and physically to confer disease resistance. *EMBO Journal* **33**, 1941–1959.
17. Zhang, L., Chen, S., Ruan, J., Wu, J., Tong, A. B., Yin, Q., Li, Y., David, L., Lu, A., LiWang, W., Marks, C., Ouyang, Q., Zhang, X., Mao, Y., and Wu, H. (2015) Cryo-EM structure of the activated NAIP2-NLRC4 inflammasome reveals nucleated polymerization. *Science* **350**, 404–409.
18. Hu, Z., Zhou, Q., Zhang, C., Fan, S., Cheng, W., Zhao, Y., Shao, F., Wang, H.-W., Sui, S.-F., and Chai, J. (2015) Structural and biochemical basis for induced self-propagation of NLRC4. *Science* **350**, 399–404.
19. Diebolder, C. A., Halff, E. F., Koster, A. J., Huizinga, E. G., Koning, R. I., Diebolder, C. A., Halff, E. F., Koster, A. J., Huizinga, E. G., and Koning, R. I. (2015) Cryoelectron tomography of the NAIP5/NLRC4 inflammasome: Implications for NLR activation. *Structure* **23**, 2349–2357.
20. Mestre, P., and Baulcombe, D. C. (2006) Elicitor-mediated oligomerization of the tobacco N disease resistance protein. *Plant Cell* **18**, 491–501.

21. Hao, W., Collier, S. M., Moffett, P., and Chai, J. (2013) Structural basis for the interaction between the potato virus X resistance protein (Rx) and its cofactor Ran GTPase-activating protein 2 (RanGAP2). *Journal of Biological Chemistry* **288**, 35868–35876.

22. Stols, L., Gu, M., Dieckman, L., Raffen, R., Collart, F. R., and Donnelly, M. I. (2002) A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expression and Purification* **25**, 8–15.

23. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expression and Purification* **41**, 207–234.

24. Anglister, J., Grzesiek, S., Ren, H., Klee, C. B., and Bax, A. (1993) Isotope-edited multidimensional NMR of calcineurin B in the presence of the non-deuterated detergent CHAPS. *Journal of Biomolecular NMR* **3**, 121–126.

25. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., Skelton, N. J., and Rance, M. (2007) *Protein NMR Spectroscopy: Principles and Practice*, 2nd ed., Elsevier

26. Fischer, H., Polikarpov, I., and Craievich, A. F. (2004) Average protein density is a molecular-weight-dependent function. *Protein Science : A Publication of the Protein Society* **13**, 2825–2828.

27. Mobli, M., Stern, A. S., Bermel, W., King, G. F., and Hoch, J. C. (2010) A non-uniformly sampled 4D HCC(CO)NH-TOCSY experiment processed using maximum entropy for rapid protein sidechain assignment. *Journal of Magnetic Resonance* **204**, 160–164.

28. Mobli, M., Maciejewski, M. W., Gryk, M. R., and Hoch, J. C. (2007) An automated tool for maximum entropy reconstruction of biomolecular NMR spectra. *Nature Methods* **4**, 467–468.

29. Mobli, M., Maciejewski, M. W., Gryk, M. R., and Hoch, J. C. (2007) Automatic maximum entropy spectral reconstruction in NMR. *Journal of Biomolecular NMR* **39**, 133–139.

30. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696.

31. Herrmann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *Journal of Molecular Biology* **319**, 209–227.

32. Guntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods in Molecular Biology* **278**, 353–378.

33. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *Journal of Biomolecular NMR* **13**, 289–302.

34. Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V., and Svergun, D. I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *Journal of Applied Crystallography* **45**, 342–350.

35. Rambo, R. P., and Tainer, J. A. (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477–481.

36. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) RIMUS : a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography* **36**, 1277–1282.

37. Semenyuk, A. V., and Svergun, D. I. (1991) GNOM – a program package for small-angle scattering data processing. *Journal of Applied Crystallography* **24**, 537–540

38. Squire, P. G., and Himmel, M. E. (1979) Hydrodynamics and protein hydration. *Archives of Biochemistry and Biophysics* **196**, 165–177.

39. Schneidman-Duhovny, D., Hammel, M., and Sali, A. (2010) FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Research* **38**, W540–544.

40. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology* **234**, 779–815.

41. Fiser, A., Do, R. K., and Sali, A. (2000) Modeling of loops in protein structures. *Protein Science* **9**, 1753–1773.

42. Newman, J. (2004) Novel buffer systems for macromolecular crystallization. *Acta Crystallographia Section D Biological Crystallography* **60**, 610–612.

43. McPhillips, T. M., McPhillips, S. E., Chiu, H. J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M., and Kuhn, P. (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *Journal of Synchrotron Radiation* **9**, 401–406.

44. Kabsch, W. (2010) XDS. *Acta Crystallographica Section D: Biological Crystallography* **66**, 125–132.

45. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the

CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography* **67**, 235–242.

46. Skubák, P., and Pannu, N. S. (2013) Automatic protein structure solution from weak X-ray data. *Nature Communications* **4**, 2777.
47. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of Biomolecular NMR* **44**, 213–223.
48. Rairdan, G. J., Collier, S. M., Sacco, M. a., Baldwin, T. T., Boettrich, T., and Moffett, P. (2008) The coiled-coil and nucleotide binding domains of the Potato Rx disease resistance protein function in pathogen recognition and signaling. *Plant Cell* **20**, 739–751.
49. Maekawa, T., Kufer, T. A., and Schulze-Lefert, P. (2011) NLR functions in plant and animal immune systems: so far and yet so close. *Nature Immunology* **12**, 817–826.
50. Bryson, K., McGuffin, L. J., Marsden, R. L., Ward, J. J., Sodhi, J. S., and Jones, D. T. (2005) Protein structure prediction servers at University College London. *Nucleic Acids Research* **33**, W36–W38.
51. Césari, S., Moore, J., Chen, C., Webb, D., Periyannan, S., Mago, R., Bernoux, M., Lagudah, E., and Dodds, P. N. (In press) Cytosolic activation of cell death and stem rust resistance by wheat MLA-family CC-NLR proteins. *Proceedings of the National Academy of Sceince U.S.A.*
52. Mago, R., Zhang, P., Vautrin, S., Simkova, H., Bansal, U., Luo, M.-C., Rouse, M., Karaoglu, H., Periyannan, S., James, K., Jin, Y., Ayliffe, M., Bariana, H., Park, R., McIntosh, R., Dolezel, J., Berges, H., Spielmeyer, W., Lagudah, E., Ellis, J., and Dodds, P. (2015) The wheat Sr50 gene reveals rich diversity at a cereal disease resistance locus. *Nature Plants* **1**, Article 15186.
53. Cesari, S., Kanzaki, H., Fujiwara, T., Bernoux, M., Chalvon, V., Kawano, Y., Shimamoto, K., Dodds, P., Terauchi, R., and Kroj, T. (2014) The NB-LRR proteins RGA4 and RGA5 interact functionally and physically to confer disease resistance. *EMBO Journal* **33**, 1941–1959.
54. Mylonas, E., and Svergun, D. I. (2007) Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *Journal of Applied Crystallography* **40**, s245–s249.
55. Wen, J., Arakawa, T., and Philo, J. S. (1996) Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Analytical Biochemistry* **240**, 155–166.
56. Liu, Y., and Eisenberg, D. (2002) 3D domain swapping: as domains continue to swap. *Protein Science* **11**, 1285–1299.
57. Bernoux, M., Ellis, J. G., and Dodds, P. N. (2011) New insights in plant immunity signaling activation. *Current Opinion in Plant Biology* **14**, 512–518.

58. Lu, A., Magupalli, V. G., Ruan, J., Yin, Q., Atianand, M. K., Vos, M. R., Schroder, G. F., Fitzgerald, K. A., Wu, H., and Egelman, E. H. (2014) Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes. *Cell* **156**, 1193–1206.

59. Lin, S. C., Lo, Y. C., and Wu, H. (2010) Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling. *Nature* **465**, 885–890.

60. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography* **66**, 12–21.

61. Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology* **292**, 195–202.

62. Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K., and Jones, D. T. (2013) Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Research* **41**, W349–357.

Chapter 5

Conclusions & Future Directions

5.1 Conclusions

5.1.1 *Project achievements overview*

The research presented in this thesis has furthered the understanding of the human innate immune signalling protein MAL and plant innate immune receptor protein Sr33. Study of these proteins has focused on structural analysis and has been summarised by a flow-diagram of achievements, which is shown in Figure 5.1.

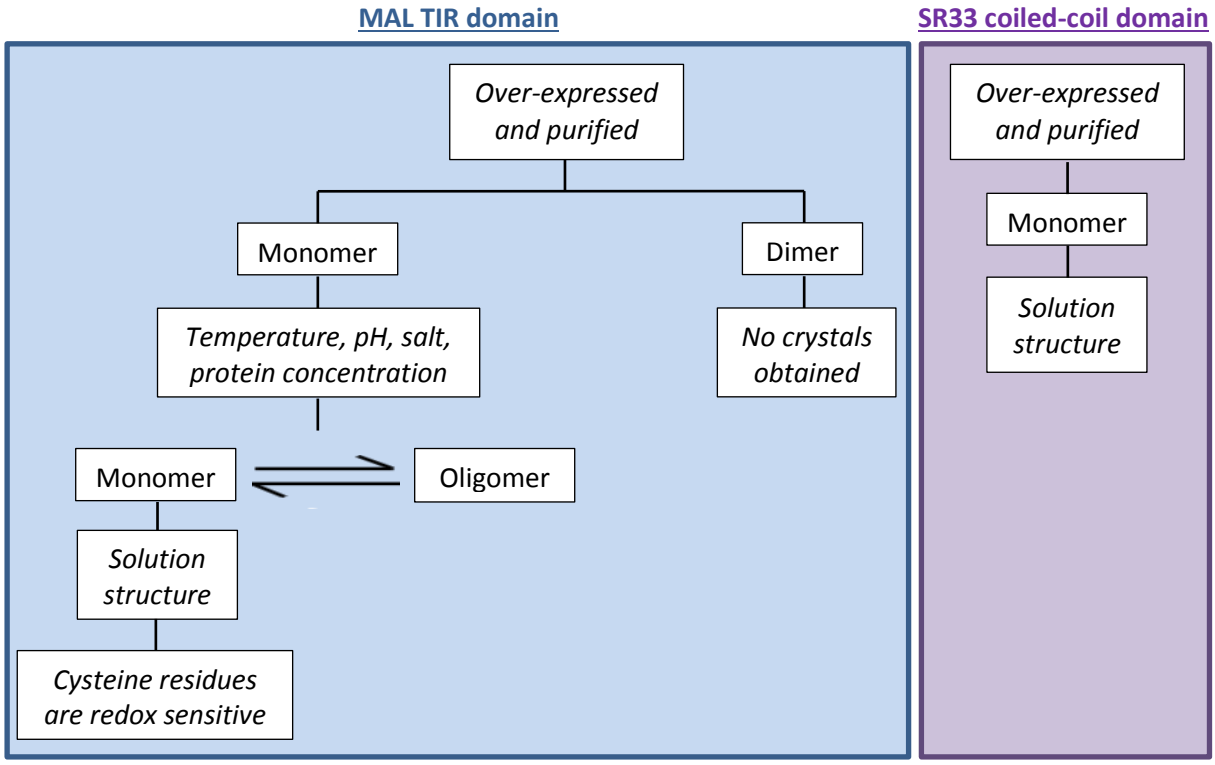


Figure 5.1. Flow diagram of project achievements for the MAL TIR domain and the Sr33 coiled-coil domain.

5.1.2 *The biochemistry and structure of the human MAL^{TIR}*

In this thesis, studies on the biochemistry of the MAL^{TIR} have shown that in solution, the protein exists in a dynamic equilibrium between a monomer and oligomer. The state of this equilibrium is dependent on a number of factors, which include: the concentration of salt, concentration of protein, temperature, and pH of the solution. This can be reversed by changing any one of these factors. For example, by lowering the temperature from 20°C to 4°C while keeping other factors constant, the equilibrium will shift to favour the monomeric state. Given that the association between TIR domains has been shown to promote signalling, as in TLRs, the identification of a

reversible oligomer formation by MAL^{TIR} suggests that this is an important feature that contributes to its functioning. During the purification, a small portion (<10%) elutes as a dimer while the rest is monomeric. This dimer is not in equilibrium with the monomer, such that if re-eluted over the size-exclusion column, the protein only elutes as a dimer. In addition, the dimer is not dependent on reduction or ionic strength for its interaction. SAXS data from this thesis has shown that the dimer which is measured does not correspond to the interfaces observed in the crystal structure of MAL^{TIR} (1). This collective evidence suggests that this may be a domain-swap dimer, whose physiological relevance is unclear.

Data presented here indicates that in solution, the structure of MAL^{TIR} is different to that of the previously published crystal structures (1-4). The $MAL^{TIRC116A}$ mutant has shown to be more stable as a monomer in solution than the wild-type, while still possessing the same fold in solution. As a result, a NMR solution structure of the $MAL^{TIRC116A}$ mutant has been solved. Overall, this solution structure contains five α-helices, which surround a five-stranded anti-parallel β-sheet, with a long BB-loop situated between the βB-strand and αB-helix. The arrangement of the β-strands and α-helices in this solution structure of $MAL^{TIRC116A}$ is consistent with a typical TIR domain (5). In context, discovery of the BB-loop in the $MAL^{TIRC116A}$ solution structure presented here is striking as this is an important feature of TIR domains and has been associated with protein interactions (6-8). In support, earlier mutation studies of residues in this region have shown a significant decrease in the activation of pro-inflammatory transcription factor NF-κB (2). This structure now provides a new template for modelling approaches and future mutational studies for functional study. Furthermore, the addition of a BB loop and lack of the AB loop in the solution structure suggests that the AB loop in the crystal structure could be the oxidised form of the protein.

Previously published crystal structures of MAL^{TIR} contained a mixture of reduced and oxidised cysteine residues with four cysteines participating in two disulfide bonds (1-4). Biochemical assays presented in this thesis showed that the redox state of cysteine residues could indeed be changed. Exposure to reducing environments such as DTT and *E. coli* BL-21 cells resulted in mostly reduced cysteine residues, while growth in *E. coli* SHuffle, Origami, and exposure to hydrogen peroxide produced mostly oxidised cysteine residues. Data from chemical shifts of all seven cysteine residues as well as intact mass spectrometry strongly suggest that the cysteine thiols are reduced in solution following purification from a reducing cell type such as *E. coli* BL-21. Given that the cytosol of a human eukaryotic cell is reducing, it is not surprising that the seven cysteines, which are all exposed to solvent, are also reduced.

Redox studies in this thesis have indicated that these cysteine residues have very different reactivates when exposed to a change in the redox condition. A physiological redox range from a reducing (-225 mV) to oxidising (-190 mV) was assessed. As expected, at -225 mV, all cysteine residues were reduced. This was consistent with the solved solution structure of $MAL^{TIRC116A}$. Strikingly, as the condition changed to an oxidising -190 mV, half of population belonging to cysteine 91 became oxidised. Given that this residue is involved in coordinating a redox sensitive molecule in the crystal structures, this result suggests that the residue is highly reactive under physiological conditions and may point to functional importance.

Mass spectrometry from this research has also demonstrated that all cysteine residues in MAL^{TIR} can be modified by glutathione; a key redox molecule in the cell cytoplasm. Redox modifications such as glutathionylation of cysteine residues in cytosolic proteins have been previously shown to be critical in modulating the functionality of proteins (9). Early work has indicated that association of glutathione transferase protein GSTO-1 with other proteins in the early TLR pathway significantly affects the downstream activation of pro-inflammatory transcription factor NF-κB (10). In addition, the redox state of the cell cytoplasm has been reported to become more oxidising upon the stimulation of the human innate immune TLR4 pathway (11,12). Given that the MAL protein is associated with the TLR4 pathway, it is likely that the cysteine residues of MAL^{TIR}, especially C91, undergo redox change, which modulates the protein's ability to signal.

5.1.3 *The structure of the Sr33 coiled-coil domain*

Two crystal structures of coiled-coil domains from two plant innate immune receptor proteins, namely Rx (13) and MLA10 (14), have shown crystal structures that are very different to each other. This thesis presents the monomeric NMR solution structure of the wheat Sr33 coiled-coil domain, which is modelled as a bundle of four α-helices. Three of these helices consistently show strong evidence for close association, while one helix (α4) at the C-terminus, is weakly associated with the remaining three helices. Given that the solution structure of the Sr33 coiled-coil domain is truncated at the C-terminus, such that the domain is no longer active as shown by Cesari *et al* (unpublished), it is possible that a longer construct which extends off the α4 helix contains important structural information. This would provide a functional reason for the weak binding α4 helix, which is currently unknown. The results presented here show that the Sr33 coiled-coil domain contains the conserved EDVID motif (EDAVD in Sr33) on the solvent accessible surface of α3 helix, which has been previously identified as an important motif for protein interactions (15). Taken together, this structure has

identified two features including, a weak binding α4 helix and the EDVID motif, that are consistent with the notion that the coiled-coil domain of Sr33 forms a signalling scaffold for downstream signalling.

5.2 Future Directions

5.2.1 *Future directed studies of the MAL^{TIR}*

In this thesis, two aspects of the MAL^{TIR} behaviour are presented. This includes the ability of cysteine residues to change redox state and the ability to form larger oligomeric structures. The future work of these aspects, which will build upon the work presented in this thesis, will be discussed, as well as preliminary results that have been obtained for the these future studies.

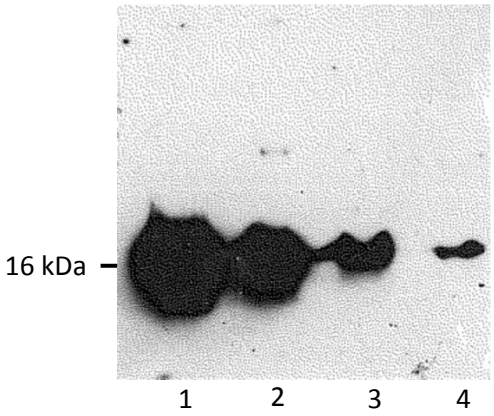
5.2.2 *In-vivo redox changes of MAL^{TIR}*

Redox modification of cysteine residues in cytosolic proteins is an important regulatory mechanism of signalling. In this study, the solution structure of the MAL^{TIRC116A} revealed seven reduced cysteine residues, which are able to undergo oxidation *in-vitro*. Further *in-vivo* studies are required to investigate the physiological relevance of the redox change on these cysteine residues. Multi-dimensional in-cell NMR experiments have shown that isotopically double-labelled protein can be used to determine chemical shifts of ^{1}H, ^{15}N and ^{13}C atoms (16). A study comparing protein insertion methods into human cell lines has found electroporation to be the most efficient method to insert labelled protein into human HEK-293T cells (17). Preliminary *in-vivo* results using this method have been successful and are presented in the following sections.

5.2.3 *MAL^{TIR} is taken up by mammalian cells*

Following expression in *E. coli* BL-21 bacterial cells, and purification using nickel affinity and size-exclusion chromatography, the wild-type myc-tagged MAL^{TIR} was mixed in a 1:1 ratio with Dulbecco's modified eagle medium (DMEM) [Thermo Fisher Scientific]. HEK-293T cells were grown in a two T7 flask until confluent (2×10^{7} cells), washed with phosphate buffered saline (PBS) solution and electroporated with the protein in a total volume of 100 µL at 250 V and 250 µ*F*. The cells were then washed to remove protein on the cell surface and rested with fresh DMEM media for four

hours. Western blot of the cell lysate following electroporation showed that the protein was taken up by the cells (lane 4 in Figure 5.2). Comparatively, two-thirds of the protein still remained in the electroporated solution without being taken-up by the cells, as shown by lane 3 in Figure 5.2.



1. Purified MAL^{TIR}
2. MAL^{TIR} with DMEM media (1:1)
3. Supernatant following electroporation
4. Cell lysate following electroporation

Figure 5.2 Western blot showing the up-take of MAL^{TIR}, which has been tagged with a Myc-tag for electroporation into mammalian HEK-293T cells. The gel shows that the protein has been taken up by the cell as shown in lane 4.

5.2.4 *Localisation of MAL^{TIR} in mammalian HEK-293T cells*

Following electroporation of the GFP-tagged MAL^{TIR} into mammalian HEK-293T cells using the protocol above, conventional wide-field and confocal fluorescence microscopy was performed to confirm up-take and localisation of the protein. Fluorescence images in Figure 5.3 show that GFP-MAL^{TIR} has been taken up by the cells. The GFP-MAL^{TIR} protein is in two forms in the cell; concentrated in a single high intensity particle, and dispersed evenly throughout the cell. Protein that has been encapsulated by an endosome reflects a higher signal intensity unit compared to cytoplasmic protein, which is evenly dispersed. Visually, the majority of the protein appears to be dispersed throughout the cell cytoplasm. In general, most of the cells have taken up equal amounts of protein, with a few cells showing high intensity, indicating a high amount of protein present. A closer examination of cellular localisation was performed using confocal microscopy, shown in Figure 5.4. Slices through HEK-293T cells show a distribution of fluorescence across the cell, indicating the GFP-MAL^{TIR} is present in the cytoplasm.

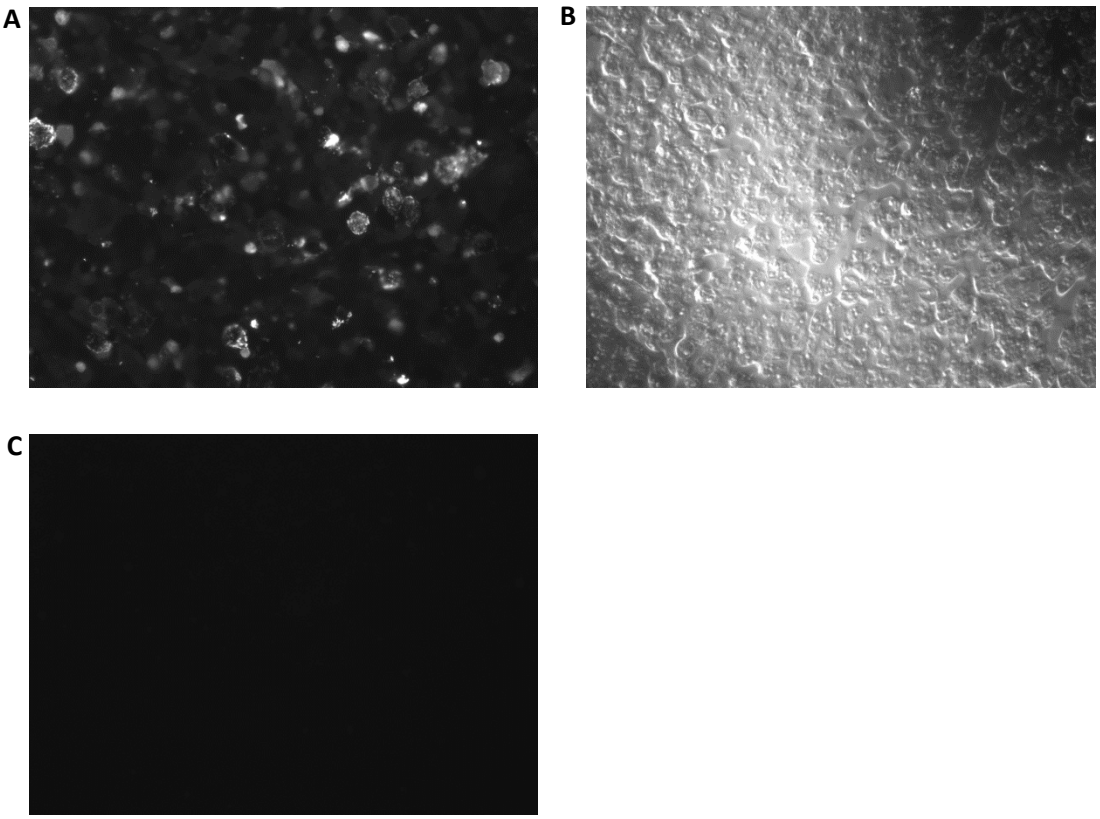


Figure 5.3 Wide-field fluorescence microscopy images obtained following electroporation with either GFP-MAL^{TIR} or Myc-MAL^{TIR} tags and one wash with PBS.

A GFP-MAL^{TIR} in HEK-293T cells

B Far view of HEK-293T cells with GFP-MAL^{TIR}

C Myc-MAL^{TIR} in HEK-293T cells (negative control)

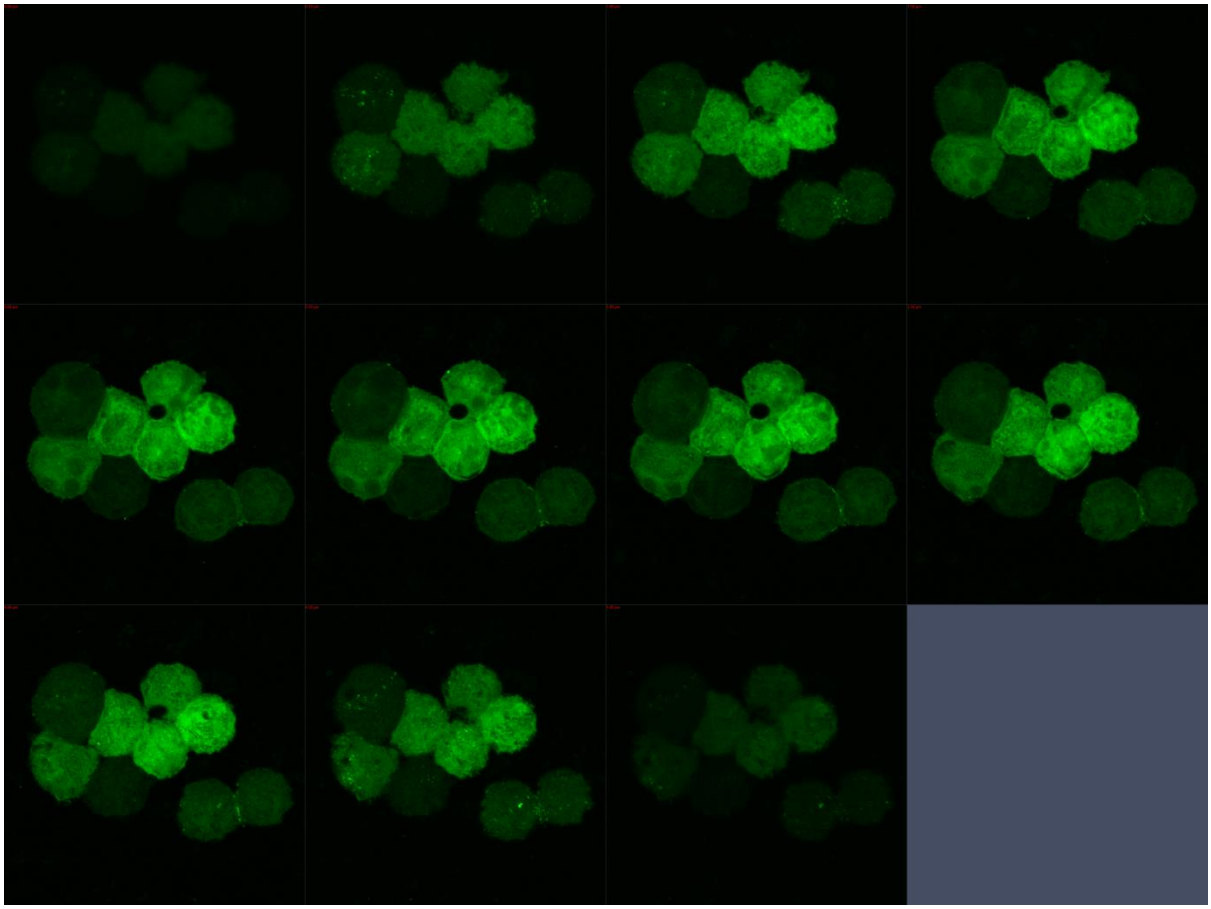


Figure 5.4 Confocal microscopy showing 0.5 µm slices through HEK-293T cells with GFP-MAL^{TIR} domain in cytoplasm. Slices through the cell show an even distribution of GFP-MAL^{TIR} throughout the cytoplasm.

5.2.5 *Future in vivo studies of MAL^{TIR}*

Preliminary results have confirmed that the protein had successfully entered the cell and the majority of the protein was indeed dispersed throughout the cytoplasm rather than in endosomes. Following this work, the next experiment would focus on confirming that the protein had indeed retained structure. It is unknown if the protein is stable in DMEM media and through the electroporation process. In addition, eukaryotic cells are continuously undergoing metabolism, which could contribute to the rapid digestion of the protein. Correct protein folding can be confirmed by ^{1}H-^{15}N-HSQC using ^{15}N-labelled protein, by comparing the spectra of the *E. coil*-produced purified protein with the electroporated intracellular protein. Further to this, it would be interesting to test the role of cysteine residues during TLR4 signalling. HEK-293T cells that contain TLR4 would be stimulated by lipopolysaccharides (LPS) and MD2, whilst containing electroporated ^{15}N and ^{13}C labelled MAL^{TIR}. By observing the β-carbon atoms belonging to the cysteine residues

(18,19), the chemical shift changes will indicate whether the redox change of the cysteine residues is involved in the signalling function of the protein.

5.2.6 *Investigation of the MAL^{TIR} oligomer*

The interaction of TIR domains has been shown to be an important feature in the signal transduction of the human immune system (20). This study has shown that the MAL^{TIR} is in equilibrium between monomeric and oligomeric states. Given that this oligomer is likely to be key to downstream signal transduction that stimulates the human immune system, there is much interest in the structure of the MAL^{TIR} oligomer as well as the mechanism that regulates the formation of the MAL^{TIR} oligomer. Un-published data by Ve *et al* has used electron microscopy to observe the structure of the MAL^{TIR} oligomer. The NMR solution structure presented in this thesis has been modelled into the density provided by electron microscopy to understand the homotypic interactions of the MAL^{TIR}. In addition, it would be interesting to investigate if the oligomer is functionally relevant. This could be tested by electroporating GFP-MAL^{TIR} into HEK-293T cells followed by induction with LPS. This would then be viewed using confocal microscopy, which would show the localisation of GFP-MAL^{TIR} from the cytoplasm to the cell membrane, followed by electron microscopy, where the oligomer would be visualised as a snap-shot of the signalling process.

5.2.7 *Future directed studies of Sr33 coiled-coil domain*

This study has determined the solution structure of the Sr33 coiled-coil domain. The structure revealed a four helix bundle. Supporting data (manuscript of Chapter 4) has defined a minimal functional unit to be 22 residues longer than that which has been solved. Future work could focus on collecting, ^{1}H, ^{15}N and ^{13}C resonances of the full Sr33 coiled-coil domain, which is functionally active, and obtain the structure to obtain functional insight. This would build on the current model presented in this thesis as the majority of the resonances could be transferred from the existing structure. Shown in Chapter 4, self-association of coiled-coil domains highly correlates with cell-death. Currently there is little information on the nature of the interaction between coiled-coil domains at a molecular level. *In-vitro* NMR experiments using a mixture of labelled and unlabelled protein could be used to determine the interacting interface of these domains. In addition, the nature of binding could be determined by assessing chemical shifts and peak intensities, providing information on the binding affinity and structural changes that occur as a result of interaction.

Further to this, the solution structure of another coiled-coil receptor protein MLA10 could be solved, which contains a crystal structure that is very different to Rx and Sr33 coiled-coil receptors domain (14). Given that the solution structure of Sr33 presented in this thesis is consistent with another receptor coiled-coil protein structure, namely Rx (13), solving the solution structure of MLA10 coiled-coil receptor which is structurally very different would begin to answer the role that the structure plays in protein interactions.

5.3 Chapter 5 references

1. Valkov, E., Stamp, A., DiMaio, F., Baker, D., Verstak, B., Roversi, P., Kellie, S., Sweet, M. J., Mansell, A., Gay, N. J., Martin, J. L., and Kobe, B. (2011) Crystal structure of Toll-like receptor adaptor MAL/TIRAP reveals the molecular basis for signal transduction and disease protection. *Proceedings of the National Academy of Sciences U S A* **108**, 14879–14884.
2. Lin, Z., Lu, J., Zhou, W., and Shen, Y. (2012) Structural insights into TIR domain specificity of the bridging adaptor Mal in TLR4 signaling. *PLoS One* **7**, e34202.
3. Snyder, G. A., Deredge, D., Waldhuber, A., Fresquez, T., Wilkins, D. Z., Smith, P. T., Durr, S., Cirl, C., Jiang, J., Jennings, W., Luchetti, T., Snyder, N., Sundberg, E. J., Wintrode, P., Miethke, T., and Xiao, T. S. (2014) Crystal structures of the Toll/Interleukin-1 receptor (TIR) domains from the Brucella protein TcpB and host adaptor TIRAP reveal mechanisms of molecular mimicry. The *Journal of Biological Chemistry* **289**, 669–679.
4. Woo, J. R., Kim, S., Shoelson, S. E., and Park, S. (2012) X-ray Crystallographic Structure of TIR-Domain from the Human TIR-Domain Containing Adaptor Protein/MyD88-Adaptor-Like Protein (TIRAP/MAL). *Bulletin of the Korean Chemical Society* **33**, 3091–3094.
5. Ve, T., Williams, S. J., and Kobe, B. (2015) Structure and function of Toll/interleukin-1 receptor/resistance protein (TIR) domains. *Apoptosis* **20**, 250–261.
6. Jiang, Z., Georgel, P., Li, C., Choe, J., Crozat, K., Rutschmann, S., Du, X., Bigby, T., Mudd, S., Sovath, S., Wilson, I. A., Olson, A., and Beutler, B. (2006) Details of Toll-like receptor:adapter interaction revealed by germ-line mutagenesis. *Proceedings of the National Academy of Sciences U S A* **103**, 10961–10966.
7. Toshchakov, V. U., Basu, S., Fenton, M. J., and Vogel, S. N. (2005) Differential involvement of BB loops of toll-IL-1 resistance (TIR) domain-containing adapter proteins in TLR4- versus TLR2-mediated signal transduction. *Journal of Immunology* **175**, 494–500.
8. Stack, J., and Bowie, A. G. (2012) Poxviral protein A46 antagonizes Toll-like receptor 4 signaling by targeting BB loop motifs in Toll-IL-1 receptor adaptor proteins to disrupt receptor:adaptor interactions. The *Journal of Biological Chemistry* **287**, 22672–22682.
9. Dalle-Donne, I., Rossi, R., Giustarini, D., Colombo, R., and Milzani, A. (2007) S-glutathionylation in protein redox regulation. *Free Radicals Biology & Medicine* **43**, 883–898.
10. Menon, D., Coll, R., O'Neill, L. A. J., and Board, P. G. (2015) GSTO1-1 modulates metabolism in macrophages activated through the LPS and TLR4 pathway. *Journal of Cell Science* **128**, 1982–1990.
11. Park, H. S., Jung, H. Y., Park, E. Y., Kim, J., Lee, W. J., and Bae, Y. S. (2004) Cutting Edge: Direct interaction of TLR4 with NAD(P)H oxidase 4 isozyme is essential for lipopolysaccharide-

induced production of reactive oxygen species and activation of NF-κB. *The Journal of Immunology* **173**, 3589–3593.

12. Ichikawa, S., Miyake, M., Fujii, R., and Konishi, Y. (2012) MyD88 associated ROS generation is crucial for Lactobacillus induced IL-12 production in macrophage. *PLoS One* **7**, e35880.
13. Hao, W., Collier, S. M., Moffett, P., and Chai, J. (2013) Structural basis for the interaction between the potato virus X resistance protein (Rx) and its cofactor Ran GTPase-activating protein 2 (RanGAP2). The *Journal of Biological Chemistry* **288**, 35868–35876.
14. Maekawa, T., Cheng, W., Spiridon, L. N., Toller, A., Lukasik, E., Saijo, Y., Liu, P., Shen, Q. H., Micluta, M. A., Somssich, I. E., Takken, F. L., Petrescu, A. J., Chai, J., and Schulze-Lefert, P. (2011) Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host & Microbe* **9**, 187–199.
15. Rairdan, G. J., Collier, S. M., Sacco, M. A., Baldwin, T. T., Boettrich, T., and Moffett, P. (2008) The coiled-coil and nucleotide binding domains of the Potato Rx disease resistance protein function in pathogen recognition and signaling. *Plant Cell* **20**, 739–751.
16. Serber, Z., Corsini, L., Durst, F., and Dotsch, V. (2005) In-cell NMR spectroscopy. *Methods in Enzymology* **394**, 17–41.
17. Bekei, B. (2013) In-cell NMR spectroscopy in mammalian cells. *Doctoral dissertation, department of biology, chemistry and pharmacy, University of Berlin*.
18. Felli, I. C., Gonnelli, L., and Pierattelli, R. (2014) In-cell (1)(3)C NMR spectroscopy for the study of intrinsically disordered proteins. *Nature Protocols* **9**, 2005–2016.
19. Banci, L., Letizia, B., Ivano, B., Enrico, L., Erica, S., Yuguang, Z., and Radu, A. A. (2014) Atomic-resolution monitoring of protein maturation in live human cells by NMR. *Nature chemical biology* **9**, 297–299.
20. Ve, T., Gay, N. J., Mansell, A., Kobe, B., and Kellie, S. (2012) Adaptors in toll-like receptor signaling and their potential as therapeutic targets. *Current Drug Targets* **13**, 1360–1374.